
The evolution of dark and luminous structure in massive early-type galaxies

By

LINDSAY OLDHAM



Institute of Astronomy
and
Corpus Christi College
UNIVERSITY OF CAMBRIDGE

Supervisors:
MATTHEW AUGER
AND
N. WYN EVANS

This dissertation is submitted for the degree of DOCTOR OF
PHILOSOPHY

MAY 2017

ABSTRACT

In this thesis, I develop and combine strong lensing and dynamical probes of the mass of early-type galaxies (ETGs) in order to improve our understanding of their dark and luminous mass structure and evolution.

Firstly, I demonstrate that the dark matter halo of our nearest brightest cluster galaxy (BCG), M87, is centrally cored relative to the predictions of dark-matter-only models, and suggest an interpretation of this result in terms of dynamical heating due to the infall of satellite galaxies. Conversely, I find that the haloes of a sample of 12 field ETGs are strongly cusped, consistent with adiabatic contraction models due to the initial infall of gas. I suggest an explanation for these differences in which the increased rate of merging and accretion experienced by ETGs in dense environments leads to increased amounts of halo heating and expansion, such that the signature of the halo's initial contraction is erased in BCGs but retained in more isolated systems.

Secondly, I find evidence that the stellar-mass-to-light ratio declines with increasing radius in both field and cluster ETGs. With M87, I show that the strength of this gradient cannot be explained by trends in stellar metallicity or age if the stellar initial mass function (IMF) is spatially uniform, but that an IMF which becomes increasing bottom-heavy towards the galaxy centre can fully reproduce the inference on the stellar mass.

Finally, I use the sizes, stellar masses and luminous structures of two samples of massive ETGs at redshift $z \sim 0.6$ to set constraints on the mechanisms of ETG growth. I find that ETGs in dense cluster environments already lie on the local size-mass relation at this redshift, contrary to their isolated counterparts, and suggest that this may be evidence for their accelerated growth at early times due to the higher incidence of merger events in clusters. I also show that massive compact ETGs at this redshift are composed of a compact, red, spheroidal core surrounded by a more extended, diffuse, bluer envelope, which may be a structural imprint of their ongoing inside-out growth. Overall, the studies presented in this thesis suggest a coherent scenario for ETG evolution which is dominated by hierarchical processes.

ACKNOWLEDGEMENTS

My first thanks are to Matt: for teaching me not only new science, but also how to do science well. Thanks for the huge amount of time you have spent with me, the freedom you have given me, and for being extremely considerate and committed to making my science meaningful and good. You have provided me with an example of how to ask questions and look for answers that I hope I will be able to reproduce one day. Also to Wyn, for being such a reliable dynamicist-on-call throughout my time here, and to Vasily and the other powers at the IoA, for unreasonably giving me a PhD position even though I didn't apply until many months too late: thanks for taking a chance on me!

To my office mates, my office-but-one mate, my year group – you've made these three years a lot of fun. To Mahsa for insistently focusing on what matters; to Rosie for the Grantchester routine; to Warda for all the joy. To the Edenfield and Theydon clans, for putting up with my hermit-like tendencies; and to my parents, for what I will summarise here as everything but particularly for being there, even (especially?) when you couldn't help. I realise a too-long thesis about a bunch of galaxies nobody's heard about might not be your top choice of reward – but it's too late for that now...

DECLARATION

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration unless explicitly specified in the text. It is not being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution, and does not exceed the limit of 60,000 words.

Lindsay Oldham

May 2017

TABLE OF CONTENTS

	Page
List of Tables	xiii
List of Figures	xiv
1 Introduction	1
1.1 Galaxy assembly in a hierarchical Universe	2
1.2 Luminous mass structure	3
1.2.1 Size evolution	3
1.2.2 The stellar initial mass function	6
1.3 Dark mass structure	8
1.3.1 The role of baryonic physics	8
1.3.2 Tools for measuring halo structure	10
1.3.3 Combining mass probes: observational results so far	12
1.4 This thesis: challenges in ETG evolution	15
2 M87's globular cluster populations	17
2.1 Introduction	17
2.2 Data & reduction	19
2.3 Photometry	20
2.4 Globular Cluster Candidates	23
2.5 Globular Cluster Populations	24
2.6 Discussion	30
2.6.1 Characterising M87's globular cluster populations	30
2.6.2 Importance of correctly modelling the interlopers	35
2.6.3 The existence of colour gradients in the globular cluster populations	38
2.7 Conclusions	39
3 M87 from parsec to megaparsec scales	43

TABLE OF CONTENTS

3.1	Introduction	44
3.2	Data	46
3.2.1	Stars	47
3.2.2	Globular clusters	48
3.2.3	Satellite galaxies	50
3.3	Modelling	52
3.3.1	Mass models	52
3.3.2	Jeans modelling	54
3.3.3	Statistical analysis	56
3.4	Results	58
3.5	Discussion	69
3.5.1	Resolving previous discrepancies	69
3.5.2	The case for a cored halo	74
3.5.3	Implications for the stellar initial mass function	76
3.5.4	Importance of correctly modelling the underlying tracer populations	77
3.5.5	Anisotropy, and a pinch radius from multiple tracer populations	78
3.6	Conclusions	79
4	Little size evolution in $z \sim 0.6$ cluster ellipticals	83
4.1	Introduction	84
4.2	Data sources and reduction	85
4.2.1	Sources	85
4.2.2	Photometry	86
4.2.3	Spectroscopy	89
4.3	Scaling relations	91
4.3.1	The Colour–Magnitude Relation	91
4.3.2	The Fundamental Plane	92
4.3.3	The $M_{dyn}/L - M_{dyn}$ relation	94
4.3.4	The $M_{\star} - \sigma$ and $M_{\star} - R_e$ relations	96
4.4	Stellar Population Models	99
4.4.1	Luminosity evolution from the CMR	101
4.4.2	Luminosity evolution from the fundamental plane	102
4.4.3	Luminosity evolution from the $M_{dyn}/L - M_{dyn}$ relation	104
4.4.4	Combining size and luminosity evolution	105
4.4.5	Combining scaling relations: inferring size and luminosity evolution	110
4.5	Discussion	112

4.5.1	Old, passively evolving stellar populations	112
4.5.2	Accelerated growth?	113
4.5.3	Can we trust the stellar and dynamical masses?	114
4.5.4	Can we compare MACSJ0717 with Coma?	114
4.6	Conclusions	115
5	Red nuggets growing inside-out	117
5.1	Introduction	118
5.2	Data	119
5.3	Lens modelling	120
5.4	Modelling Results	128
5.4.1	Accurately modelling the EELs	128
5.5	Source galaxy properties	130
5.5.1	Stellar masses	133
5.5.2	The observed size-mass relation	134
5.5.3	The intrinsic size-mass relation	136
5.5.4	Morphologies	137
5.6	Discussion	139
5.6.1	Are the EELs sources red nugget relics?	139
5.6.2	Are the EELs sources evolving red nuggets?	140
5.6.3	Colour gradients and inside-out growth	141
5.6.4	Growth in dense environments?	143
5.7	Summary and conclusions	144
6	The fundamental plane of evolving red nuggets	147
6.1	Introduction	148
6.2	Data and kinematic modelling	149
6.3	Fundamental plane	153
6.3.1	The observed fundamental plane	153
6.3.2	The intrinsic fundamental plane	156
6.4	Physical models	158
6.5	Discussion	160
6.5.1	The fundamental plane	160
6.5.2	Physical models	164
6.6	Conclusions	166
7	M87's radially varying IMF	167

TABLE OF CONTENTS

7.1	Introduction	168
7.2	Data	170
7.2.1	Photometry	170
7.2.2	Kinematics	171
7.3	Dynamical model	173
7.3.1	Mass components	173
7.3.2	Stellar mass-to-light ratio	173
7.3.3	Anisotropy	174
7.3.4	Large-radius mass	175
7.4	Statistical model	175
7.5	Results	176
7.6	Stellar population modelling	178
7.7	Discussion	185
7.7.1	Inference on the halo structure is robust	185
7.7.2	Radial gradients in the IMF?	185
7.7.3	The importance of accounting for IMF gradients	187
7.8	Summary and conclusions	188
8	Dark and luminous mass structure of strong lenses	189
8.1	Introduction	190
8.2	Data and modelling	191
8.2.1	Data	192
8.2.2	Mass model	192
8.2.3	Lens modelling	193
8.2.4	Dynamical modelling	194
8.3	Results	196
8.3.1	Models with a spatially constant Y_{\star}	196
8.3.2	Models with a spatially varying Y_{\star}	199
8.3.3	The population as a whole	205
8.3.4	Systematics	206
8.4	Discussion	210
8.4.1	Dark matter	210
8.4.2	Stellar mass and the IMF	212
8.4.3	The evolution of massive ETGs	215
8.5	Summary and conclusions	216

9	Conclusions	217
9.1	What is the structure of the dark halo in ETGs?	217
9.2	What is the nature of the IMF in ETGs?	218
9.3	What are the physical mechanisms driving ETG growth?	219
9.4	A coherent view of ETG evolution	220
9.5	Future directions	220
A	Anisotropy kernel for the scaled Osipkov-Merritt profile	223
B	Individual MACSJ0717 galaxies	225
C	Notes on individual EELs	229
D	Pixellated reconstructions of the EELs	233
E	Inference on M87's mass structure	237
	Bibliography	243

LIST OF TABLES

TABLE	Page
2.1 Globular cluster candidate catalogue	27
2.2 Inference on globular cluster spatial distributions	33
3.1 Sources of data for dynamical modelling of M87	46
3.2 Inference on isotropic dynamical models of M87	58
3.3 Inference on anisotropic dynamical models of M87	71
4.1 Photometry and kinematics of MACSJ0717 galaxies	90
4.2 Inference on scaling relations for MACSJ0717 galaxies	107
4.3 Inference on scaling relations for Coma galaxies	108
4.4 Inference on structural and stellar population properties from combined scaling relations	109
5.1 Locations and redshifts of the EELs	120
5.2 Lens models for the EELs	127
5.3 Source models for the EELs	129
5.4 Stellar masses for the EELs	130
5.5 Size-mass relation of the EELs source galaxies	133
6.1 Fundamental plane data for the EELs source galaxies	153
6.2 Fundamental plane of the EELs source galaxies	155
6.3 Dark halo structure of the EELs source galaxies	160
7.1 Inference on M87's mass allowing a varying stellar-mass-to-light ratio	181
8.1 Constant stellar-mass-to-light models for the EELs lenses	203
8.2 Varying stellar-mass-to-light ratio models for the EELs lenses	203
8.3 Hierarchical models of the EELs as a population	205
8.4 Correlations between EELs mass properties	205

LIST OF FIGURES

FIGURE	Page
1.1 ETG stellar-mass-size relation	4
1.2 The stellar initial mass function	7
1.3 Strong lensing geometry	12
1.4 Combining strong lensing and kinematics	13
2.1 Step-by-step light subtraction from M87	21
2.2 Point source selection with SExtractor	25
2.3 Colour selection of globular cluster candidates	26
2.4 Luminosity function of interloping populations	30
2.5 Inference on globular cluster spatial distributions	31
2.6 Inferred glbular cluster surface density	32
2.7 Contamination from Sagittarius stars	37
2.8 Histograms of radially varying globular cluster colour distributions	38
2.9 Globular cluster colour distributions	40
3.1 Density profiles of M87's kinematic tracers	46
3.2 M87's stellar surface density profile	48
3.3 Spatial distributions of M87's kinematic tracers	49
3.4 Robustness of halo mass estimator	52
3.5 Inference on isotropic NFW model	59
3.6 Inference on isotropic LOG model	60
3.7 Inference on isotropic gNFW model	61
3.8 Inference on isotropic cgNFW model	62
3.9 Comparison of halo models	63
3.10 Inference on anisotropic NFW model	64
3.11 Inference on anisotropic LOG model	65
3.12 Inference on anisotropic gNFW model	66

3.13	Inference on anisotropic cgNFW model	67
3.14	Comparison of model and data for M87's kinematics (SAURON)	69
3.15	Inference on M87's mass profile	70
3.16	Inference excluding stellar kinematics	73
3.17	Inference on logarithmic halo density slope	76
3.18	Evidence for a pinch radius	80
4.1	HST image of MACSJ0717	87
4.2	Colour-magnitude diagram of MACSJ0717 galaxies	91
4.3	Fundamental plane of MACSJ0717 and Coma galaxies	93
4.4	MLM relation for MACSJ0717 and Coma galaxies	96
4.5	Mass-size and mass-velocity-dispersion relations	100
4.6	Constraints on the stellar age	103
5.1	Data and lens models for the EELs	121
5.2	Observed size-mass relation of the EELs source galaxies	131
5.3	Intrinsic size-mass relation for the EELs source galaxies	132
5.4	Colour gradients in the EELs source galaxies	143
6.1	Keck/ESI spectra for the EELs	151
6.2	Observed fundamental plane of the EELs source galaxies	157
6.3	Dark halo structure of the EELs source galaxies	161
6.4	Mass-luminosity relation of the EELs source galaxies	164
7.1	Luminosity profiles of M87's kinematic tracers (2)	171
7.2	Comparison of model and data for M87's kinematics (MUSE)	177
7.3	Inferred mass profile for M87	179
7.4	Stellar-mass-to-light ratio profile of M87	180
7.5	Inference on M87's IMF slope	183
7.6	Inference on M87's logarithmic halo density slope	184
8.1	Inference on dark and light lens models for a typical EEL	195
8.2	Reconstructed image for a typical EEL	196
8.3	Inference on dark and light mass structure for the EELs	200
8.4	Correlations of EELs mass properties	201
8.5	Systematics in the construction of lens models for the EELs	204
8.6	Evidence for an environmentally-dependent halo structure	209

B.1	Imaging and spectra for MACSJ0717 galaxies	228
D.1	Pixellated source reconstructions of the EELs	236
E.1	Inference on isotropic PL model of M87's mass	237
E.2	Inference on anisotropic PL model of M87's mass	238
E.3	Inference on isotropic SC model of M87's mass	239
E.4	Inference on anisotropic SC model of M87's mass	240
E.5	Inference on M87's stellar population properties	241

CHAPTER 1

INTRODUCTION

Galaxies in the Universe span a huge visual diversity. Starting with Hubble’s ‘tuning fork’ diagram (Hubble, 1926, 1927), which first classified galaxies according to their appearance on the sky, it has become clear that the structure of a galaxy retains much information about its assembly and evolution. This fact has been key in allowing us to (a) simplify the ‘zoo’ of galaxies that we observe and (b) access the underlying physical processes that govern their growth. One major insight from these efforts is that the diversity we see can be reformulated, to some extent, as a sequence, and that galaxies evolve *hierarchically* from young, low-mass, star-forming, disk types into passive, massive, spheroidal ones (e.g. Larson, 1990). However, the mechanics of this evolutionary process remain difficult to understand, and new data-driven discoveries such as the compactness of high-redshift early-type galaxies (ETGs) and the existence of ultra-diffuse galaxies are now challenging us to update and improve this simple theoretical picture. Moreover, the discovery that galaxies reside in haloes of dark matter which hugely dominate the baryonic mass requires us to understand how these two structural components co-evolve. This thesis is concerned with disentangling the dark and luminous mass and so tracing the action of these physical processes in ETGs, which, as the most massive and oldest systems in the Universe, have potentially the most to tell us about how galaxies grow.

1.1 Galaxy assembly in a hierarchical Universe

In the concordance cosmological model, the Universe is expanding from its original extremely hot, dense state following a Big Bang, and has 85% of its mass in cold, collisionless, non-radiative dark matter which interacts (possibly solely) via gravity (White & Rees, 1978). The formation of structure is then driven by the clumping of dark matter, which collapses gravitationally along filaments to form haloes which themselves merge hierarchically. It is in this cosmological context, through the infall of baryonic material onto these high-density haloes, that luminous galaxies form.

However, whilst the cosmological model has been extremely successful on large scales, it cannot provide a complete prescription for galaxy assembly on scales where baryonic mass dominates, due to the action of additional astrophysical processes such as star formation, winds from stellar evolutionary processes and feedback from active galactic nuclei (AGN). An understanding of how these processes modulate the larger-scale, cosmologically-driven evolution is therefore essential for a complete picture of galaxy evolution.

Historically, our understanding of galaxies on scales where baryonic matter dominates has been driven by the discovery of the tight scaling relations that exist between their global properties. For ETGs,¹ this began with the Faber-Jackson relation (the correlation between total galaxy luminosity and central stellar velocity dispersion; Faber & Jackson, 1976) and the Kormendy relation (the anticorrelation of effective radius with central surface brightness; Kormendy, 1977), which were later discovered to be projections of the fundamental plane (FP; Djorgovski & Davis, 1987). The FP relates the effective radius R_e , stellar velocity dispersion σ_* and surface brightness at the effective radius $\log I_e$ as $\log R_e = \alpha \log \sigma_* + \beta \log I_e + C$ with $\alpha \sim 1.2$, $\beta \sim -0.8$, and can be understood physically as a result of ETGs being homologous, virialised systems with total mass-to-light ratios Y_{tot} which increase with stellar mass. Since its discovery, the FP has been shown to have an extremely small intrinsic scatter (Jorgensen et al., 1996; Hyde & Bernardi, 2009; Graves & Faber, 2010) and to evolve negligibly with redshift, such that its evolution is consistent with the passive fading of old stellar populations; additionally, the replacement of surface brightness I_e with the surface stellar mass density Σ_* using strong lens systems has allowed the construction of the mass-FP, for which the scatter is yet smaller (Bolton et al., 2007; Auger et al., 2009).

The existence of the FP implies that all ETGs must form and evolve via processes which are almost entirely specified by their size, stellar mass and luminosity, regardless

¹Note that, in this thesis, we are focusing almost exclusively on slow-rotating ETGs, which Cappellari et al. (20116) identifies as dominating the galaxy population at masses above $\sim 2 \times 10^{11} M_\odot$.

of the properties of their stellar populations (e.g. star formation history, stellar initial mass function, metallicity and age), dark halo (e.g. mass and concentration), environment and merger history. Moreover, the discovery of the tight relation between stellar velocity dispersion and central black hole mass (Ferrarese & Merritt, 2000; McConnell & Ma, 2013) means that the evolution of the black hole can also be specified by these same few global properties. Whether similar relations exist between dark and luminous properties – such that the dark halo structure can also be determined from these few observables – remains to be seen, though the success of abundance matching techniques at reproducing galaxy clustering statistics suggests that this might be the case (Behroozi et al., 2010). The task is now to unite this phenomenological picture, which describes the bright, central regions of ETGs, with the cosmological one which operates on large scales, in order to develop a coherent scenario for how these systems evolve.

1.2 Luminous mass structure

The most observationally efficient way to probe galaxy structure is photometrically. Though dark matter contributes the majority of the total mass budget of a galaxy, the centres of massive ETGs are completely baryon-dominated, which means that their light distributions contain information about both the assembly history of the galaxy and the central mass structure. Though the FP shows no significant evolution with redshift, its projection in size-stellar mass space (the $R_e - M_\star$ relation; Shen et al., 2003; Hyde & Bernardi, 2009) exhibits strong evolution out to $z \sim 2$, such that ETGs at those redshifts have masses comparable to the local population but are 3-5 times smaller (Daddi et al., 2005; Trujillo et al., 2007; van Dokkum et al., 2008, and Figure 1.1). The implication of this is that ETGs have experienced dramatic size growth over the past 10 Gyr.

1.2.1 Size evolution

The evolution of the $R_e - M_\star$ relation for ETGs is surprising as it implies that these systems must grow in size at approximately constant stellar mass, and so requires a growth mechanism that predominantly *redistributes* mass, rather than adds to it. Furthermore, constraints from the FP and other stellar population diagnostics indicate that the stellar populations are passive, and thus rule out growth by the formation of young stars at large radii or the accretion of such stars in gas-rich (‘wet’) mergers (e.g. Treu et al., 2005). Equally, gas-poor (‘dry’) mergers between two galaxies of similar initial masses (‘major’ mergers) cause an increase in mass which scales linearly with the increase in size (e.g.

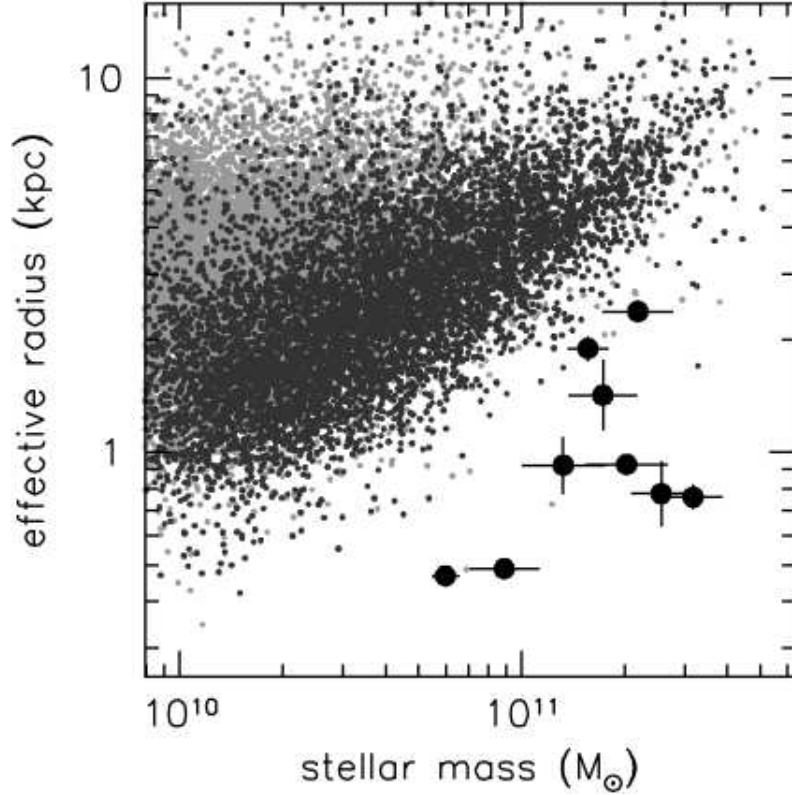


Figure 1.1: The stellar-mass-size relation for $z \sim 2.3$ ETGs compared to $z \sim 0$ SDSS galaxies. Black and grey points show red-sequence and non-red-sequence SDSS galaxies; large black points are the spectroscopically-confirmed sample of ETGs at $z \sim 2.3$ of van Dokkum et al. (2008), from which this image is reproduced. The $z \sim 2.3$ ETGs occupy the high-mass end of the local size-mass relation, but are ~ 4 times more compact.

Bezanson et al., 2009), and so would produce an excess of high-mass galaxies compared to the population we see locally.

There remain, then, two candidate mechanisms which may be able to reproduce the required size growth without giving rise to discrepancies with other local observations. Firstly, size growth may be coupled to the action of winds driven by supernovae or AGN (Fan et al., 2008; Damjanov et al., 2009; Fan et al., 2010); in this case, large gas outflows lead to a reduction of the central gravitational potential, to which the stellar content of the galaxy responds by expanding adiabatically. However, since this mechanism depends on the stellar populations being young and the galaxy’s AGN being active, it is not clear that it can produce the sustained and gradual growth that the evolution of the $R_e - M_\star$ relation implies.

On the other hand, dry mergers in which massive galaxies accrete much lower-mass ones (‘minor’ mergers) can cause efficient size growth in which the effective radius increases as the square of the mass (e.g. Bezanson et al., 2009), such that their repeated action can significantly increase the size without resulting in excessively high stellar masses. Indeed, minor mergers (both wet and dry) are expected in a hierarchical Universe, and observational evidence for such events exists in the form, for instance, of closely interacting galaxy pairs and galaxies with extended tidal tails (e.g. Elmegreen et al., 1998; Struck & Smith, 2012; Atkinson et al., 2013). As a result, dry minor mergers are currently thought to be the dominant channel for ETG size growth.

However, a number of problems remain with this explanation, and indicate the need for the development of both new observational tests and more sophisticated theoretical models. Firstly, whilst the evolution in the $R_e - M_\star$ relation becomes steeper beyond $z > 1$ – implying higher merger rates at this epoch – deep photometric censuses indicate that the abundance of satellite galaxies around compact ETGs is constant with redshift. Thus there appear to be insufficient numbers of satellites at $1 < z < 2$ to give rise to the high merger rates that the evolution in the $R_e - M_\star$ relation requires (Newman et al., 2012). Whilst this discrepancy may be partly alleviated by improvements in theoretical merger models or in our understanding of the incompleteness of deep photometric surveys, the implication is that the minor merger paradigm cannot be a complete explanation for the size evolution that we observe at these redshifts.

Furthermore, purely dry minor merger models cannot simultaneously reproduce the observed evolution of the *luminous* structure (from the $R_e - M_\star$ relation) and that of the *mass* structure (which is constrained via strong gravitational lensing and quantified in terms of the logarithmic slope of the total mass density), even below redshifts $z < 1$ (Sonnenfeld et al., 2014). Though more sophisticated theoretical minor merger models may eliminate some of this tension, Sonnenfeld et al. (2012) also showed that the addition of a $\sim 8\%$ gas fraction – i.e. a small amount of wet merging, leading to star formation – can alleviate the discrepancy to within 2σ . However, it is unclear whether the star formation rates implied by such a model are consistent with ETG spectra.

At present, it is debatable whether the introduction of additional growth processes is the correct solution to these inconsistencies. That is, the $R_e - M_\star$ evolution that we observe may be due to not only the evolution of *individual* galaxies, but also the evolution of the *population*, if the ETGs that are added to the population at later times are formed from generally more massive progenitors (that is, the progenitor population also evolves with time; e.g. van Dokkum & Franx, 2001). This *progenitor bias* hypothesis can be tested by tracing the evolution of the number density of ETGs as a function of size, but studies

so far have been inconclusive (e.g. van der Wel et al., 2014; Damjanov et al., 2015). One possibility for making progress in disentangling these effects, which we consider in this thesis, is to turn our attention from the population of ETGs as a whole to its individual members, and probe these systems structurally at high resolution in order to look for new, complementary imprints of their growth.

1.2.2 The stellar initial mass function

A further complication in understanding ETG stellar mass structure is that the conversion of light to mass requires knowledge of the stellar initial mass function (IMF), which describes the distribution of masses with which a galaxy forms stars. Though the IMF may depend on the physical conditions of the star-forming regions (in particular, the Jeans mass is smaller in cooler, denser environments, such that lower-mass stars can be formed; see McKee & Ostriker, 2007, for a review), resolved star counts have found that a spatially uniform or ‘universal’ IMF is sufficient to describe the stellar populations across the diversity of the Milky Way’s environments (Bastian et al., 2010), and that the IMF can be well parameterised as a three-segment power law, such that the relative fraction of low-mass to high-mass stars decreases towards lower masses (Kroupa, 2001, see also Figure 1.2). As a result of this apparent universality, it is common practice to assume that a *Milky-Way-like* IMF is also an appropriate description of the stellar populations in external galaxies; indeed, the IMF is an essential ingredient for the extraction of many stellar population properties extragalactically, and the assumption of an incorrect IMF could systematically bias measurements of a large number of galaxy properties.

Nevertheless, recent constraints from strong lensing and dynamics have found that the IMF in massive ETGs may be significantly *heavier* than the Milky Way’s, resulting in *larger* stellar mass-to-light ratios for any given set of stellar population properties; moreover, there is tentative evidence that the IMF becomes increasingly heavy with galaxy mass (Auger et al., 2010b; Cappellari et al., 2012). At the same time, independent constraints from stellar population modelling have found evidence for excess absorption in gravity-sensitive features in ETG spectra, indicating the presence of increased fractions of low-mass stars relative to the Milky Way (van Dokkum & Conroy, 2010), and a potential trend between the dwarf-star fraction and metallicity. Taken together, these results imply that the IMF in massive ETGs is more *bottom-heavy* than in our own Galaxy, and can be better described by a single power law which behaves similarly to the Milky Way’s IMF at high masses but does not exhibit a turnover at the low-mass end (Salpeter, 1955, and Figure 1.2); however, the driving property (e.g. galaxy mass or metallicity) behind

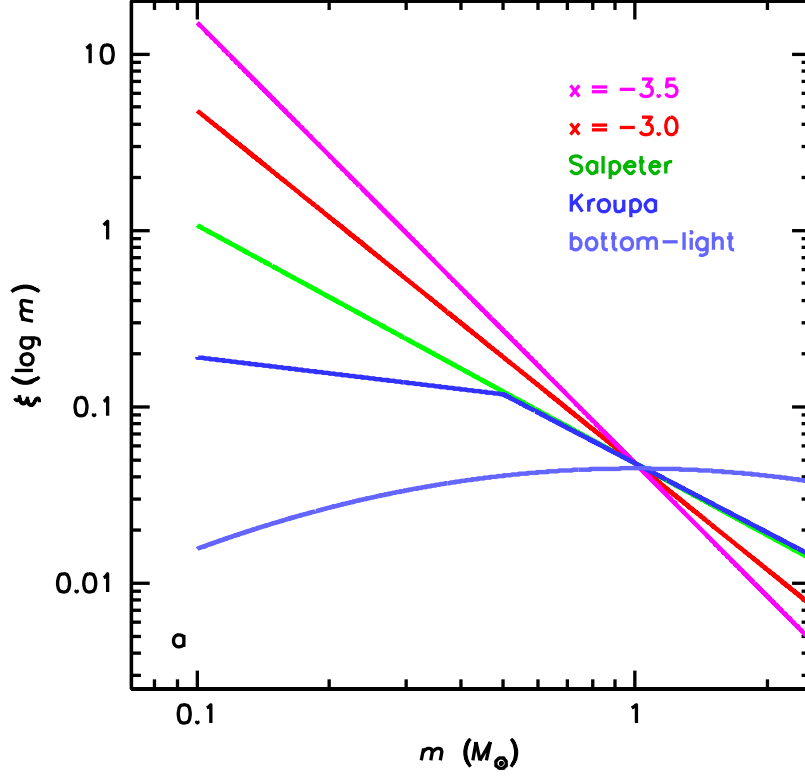


Figure 1.2: Models for the IMF. The diverse environments within the Milky Way are consistent with having a Kroupa IMF (dark blue line), whereas evidence from stellar population modelling, strong lensing and dynamics suggests that massive ETGs have higher fractions of low-mass stars, and therefore bottom-heavy IMFs that are more consistent with the Salpeter parameterisation (green line). Alternative forms for the IMF with lower or higher fractions of low-mass stars are described as ‘bottom-light’ (light blue line) or ‘bottom-heavy’ (red and pink lines) respectively, with the latter characterised by a logarithmic slope $\frac{d \log \xi}{d \log M} = x$. This figure is reproduced from van Dokkum & Conroy (2010).

these variations remains unclear. Furthermore, very recent stellar population studies have found evidence for IMF variations *within individual ETGs*, such that the IMF is bottom-heavy centrally but becomes Milky-Way-like at larger radii (Martín-Navarro et al., 2015a; van Dokkum et al., 2016, though see also Davis & McDermid, 2017 for a study based on molecular gas kinematics which found no significant trends). Currently, then, a number of open questions remain about the form of the IMF – and hence the stellar mass and stellar population properties – which must be answered by the development and application of robust, independent probes of the stellar mass to larger galaxy samples.

The extraction of the stellar mass of an ETG is further complicated, however, by the fact that this baryonic material is contained within a massive halo of dark matter which also contributes to the total mass; therefore, robustly determining the stellar mass relies on an accurate separation of the total mass into its dark and luminous components. This is a non-trivial undertaking, and is the topic of the next section.

1.3 Dark mass structure

In a Λ CDM Universe, the visible parts of galaxies are surrounded by massive dark matter haloes which have collapsed under the force of gravity, and dynamical studies have shown conclusively that the dark matter dominates the total galaxy mass across all galaxy types. It is therefore not sufficient to understand ETG structure and evolution in terms of the luminous mass alone. In particular, whilst it was initially thought that inferences on the dark halo structure would allow galaxy-scale tests of Λ CDM, it is now clear that this *cosmological* signal is completely overwhelmed by the action of the *baryonic processes* which occur during galaxy assembly and evolution. Probes of the dark halo structure are therefore highly complementary to those which focus only on the luminous mass.

1.3.1 The role of baryonic physics

In a Universe containing only dark matter and no baryons, cosmological simulations predict that all dark matter haloes should look nearly self-similar, regardless of their mass scale, such that their mass densities can be well described as following a double power law which declines as r^{-3} at large radii and asymptotes to a central r^{-1} cusp. This can be modelled using the simple two-parameter Navarro-Frenk-White (NFW) profile:

$$(1.1) \quad \rho(r) = \frac{\rho_c \delta_c}{(r/r_s)(1 + r/r_s)^2}$$

for density ρ , radius r , scale radius r_s , critical density ρ_c and characteristic halo density δ_c (Navarro et al., 1997; Navarro et al., 2010). Moreover, the strong correlation between halo virial mass M_{200} (formally, the halo mass within the radius in which the mean density is equal to $200\rho_c$) and δ_c means that such haloes can be specified by a single parameter, though with significant scatter (e.g. Macciò et al., 2008). This appears to be a general consequence of the gravitational collapse of cold, collisionless material, with no dependence on the power spectrum of the primordial density fluctuations or the cosmological parameters.

However, the presence and evolution of baryonic material almost certainly complicates this simple theoretical picture, as it may irreversibly modify the gravitational potential well of the halo; this is especially important in the centres of ETGs, where the baryonic mass is dominant. On the one hand, the initial infall of radiative gas onto a pristine halo will deepen the potential well and so contract the halo; if this process is sufficiently slow as to proceed adiabatically, then standard assumptions about the stellar mass distribution and orbital structure predict that the halo should become significantly cuspier than the NFW profile centrally, scaling as $r^{-\gamma}$ with $\gamma \sim 1.4 - 1.8$ (Blumenthal et al., 1986; Gnedin et al., 2004), and more sophisticated treatments of this contraction using hydrodynamical simulations confirm this result (Duffy et al., 2010).

At later times, though, if ETGs grow by accreting many smaller systems – which are already in the form of dense, lumpy galaxies as opposed to diffuse clouds of gas – then the orbital decay of these systems by dynamical friction will transfer energy and angular momentum to the halo and so cause it to expand (El-Zant et al., 2001, 2004; Nipoti et al., 2004). If the effect of this is significant, it may wash out some of the initial contraction, and so reduce the central halo slope by $\Delta\gamma \sim 0.3 - 0.5$ (Laporte et al., 2012). It is also possible that repeated bursts of AGN-driven gas outflows may irreversibly alter the gravitational potential such that the distributions of both the dark and luminous mass become centrally cored, though the importance of this process depends on the strength of the AGN feedback itself, which is not well constrained (Martizzi et al., 2012; Martizzi et al., 2013).

The inner structure of the halo therefore contains a large amount of information about the relative importance of different baryonic processes in the assembly and evolution of ETGs. In light of this, it is useful to define a new, three-parameter model for the halo, the so-called generalised NFW (‘gNFW’) profile:

$$(1.2) \quad \rho(r) = \frac{\rho_c \delta_c}{(r/r_s)^\gamma (1 + r/r_s)^{3-\gamma}}$$

which has the same behaviour as the original NFW profile at large radii but an inner slope γ which may deviate from unity in response to the action of baryonic processes. Meaningful inference on γ thus becomes a key goal, and is a major topic of this thesis.

1.3.2 Tools for measuring halo structure

Since most mass probes accessible with ETGs are only sensitive to the *total* mass structure, identifying the relative importance of distinct astrophysical processes in the centres of ETGs requires the robust disentanglement of the dark matter from the baryonic mass. However, the dominance of baryonic material in the central regions (dark matter makes up typically $\sim 13\%$ of the total mass within the effective radius; Cappellari et al., 2013a), together with the unknown IMF and inner halo slope, makes this difficult, such that, given a single probe of the total mass, these two components are extremely degenerate. On the other hand, ETGs are large, massive systems, for some of which it is possible to obtain and combine multiple independent mass probes, each sensitive to the total mass at different radii or in different projections. In the context of well-motivated models, these multiple measurements make it possible to break this dark/light degeneracy and so infer both the stellar mass-to-light ratio – and hence the IMF – and the inner dark matter structure simultaneously. In particular, strong gravitational lensing and dynamics are extremely useful and complementary probes of ETG mass.

1.3.2.1 Strong gravitational lensing

In General Relativity, the curvature of spacetime near massive objects causes the paths of photons to be perturbed, such that light from a distant galaxy passing close to the centre of an intervening mass is deflected (Einstein, 1936). If the foreground object has a sufficiently high surface mass density, then it acts as a *strong gravitational lens* of the background source, which appears magnified and multiply imaged on the sky. With strong lensing it is therefore possible to super-resolve and magnify faint, distant galaxies which would otherwise be difficult to observe in any detail. It also allows an extremely robust measure of the projected mass distribution of the lensing galaxy, which can be constrained to within a few percent accuracy with no assumptions about its dynamical state.

On the scale of galaxy-galaxy lenses, the cosmological distances between the Earth (the observer), the lens and the source are sufficiently large relative to the extent of the lens itself that the latter can be treated as a thin sheet (the thin-lens approximation), in which case the positions of the background galaxy in the source plane (the *true* position, $\vec{\beta}$) and the image plane (the *observed* position, $\vec{\theta}$) are related by the lens equation

$$(1.3) \quad \vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta})$$

where $\vec{\alpha}(\vec{\theta})$ is the reduced deflection angle, which can be calculated from the convergence

$\kappa(\vec{\theta})$ as

$$(1.4) \quad \tilde{\alpha}(\vec{\theta}) = \frac{1}{\pi} \int d^2\vec{\theta}' \frac{(\vec{\theta} - \vec{\theta}')\kappa(\vec{\theta}')}{|\vec{\theta} - \vec{\theta}'|^2}$$

where the convergence $\kappa(\vec{\theta})$ is the scaled surface mass density $\Sigma(D_d\vec{\theta})$

$$(1.5) \quad \kappa(\vec{\theta}) = \frac{\Sigma(D_d\vec{\theta})}{\Sigma_{cr}}$$

and the critical surface density for lensing is

$$(1.6) \quad \Sigma_{cr} = \frac{c^2 D_s}{4\pi G D_{ds} D_d}$$

with distances defined in Figure 3. The condition for strong lensing is that the surface mass density is greater than or equal to the critical surface density where the light passes the lens. (For a thorough presentation of the strong lensing formalism, see Schneider et al., 1992.) It is therefore possible to use galaxy-scale lenses with images which form extended arcs to (a) delens the background source to study its structure and (b) probe the projected surface mass density around the Einstein radius in detail.

1.3.2.2 Dynamics

A galaxy can be modelled as a collisionless system in which the total population of a particular tracer particle (e.g. stars) is conserved. The tracers then obey the collisionless Boltzmann equation (see e.g. Binney & Tremaine, 2008), the first moments of which relate the velocity dispersion $\sigma(r)$ of the tracer particles to the mass distribution $M(r)$ in which they are moving. These are the second-order Jeans equations, which in spherical coordinates reduce to

$$(1.7) \quad \frac{d}{dr}(l\sigma_r^2) + 2\frac{\beta(r)}{r}l\sigma_r^2 = l(r)\frac{GM(r)}{r^2}$$

where $l(r)$ is the luminosity density of the tracers, $\beta(r) = 1 - \sigma_t^2/\sigma_r^2$ is the anisotropy parameter and $\sigma_r(r)$, $\sigma_t(r)$ the radial and tangential velocity dispersions respectively. Thus, assuming the luminosity density is known (i.e. via a deprojection of the surface brightness profile), it is possible to infer the mass and anisotropy distribution by calculating the radial velocity dispersion and projecting it to give the line-of-sight velocity dispersion $\sigma_{los}(R)$ as a function of projected radius R

$$(1.8) \quad \frac{1}{2}I(R)\sigma_{los}(R)^2 = \int_R^\infty \frac{l\sigma_r^2 r dr}{\sqrt{r^2 - R^2}} - R^2 \int_R^\infty \frac{\beta l\sigma_r^2 dr}{r\sqrt{r^2 - R^2}}$$

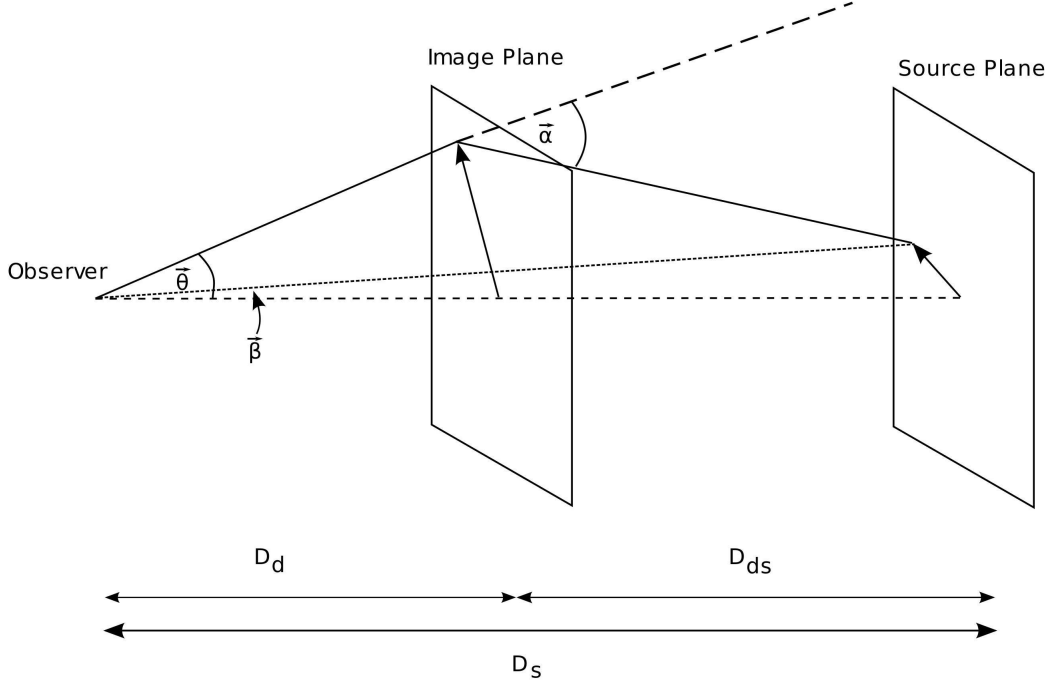


Figure 1.3: The geometry of a strong lensing system under the thin lens approximation; the lens is located in the image plane, and distances represent angular diameter distances. This diagram is reproduced from Treu (2010) and was originally created by B. Brewer.

(e.g. Mamon & Łokas, 2005), which can then be compared with observations. This approach suffers from an intrinsic degeneracy between mass and anisotropy; however, the combination of Equation 1.8 with higher-order Jeans equations, or the combination of multiple independent dynamical tracers such as stars, globular clusters, planetary nebulae and satellite galaxies – each of which has a different luminosity distribution but moves in the same gravitational potential – makes it possible to break this degeneracy and so infer the mass across a large radius range.

1.3.3 Combining mass probes: observational results so far

Strong lensing and dynamics are highly complementary mass probes since the former measures the projected mass within the Einstein radius whereas the latter is sensitive to the mass enclosed within a 3D aperture (Figure 1.4). The combination of these two independent probes therefore allows detailed inference on an ETG’s mass structure. This concept was initially applied to samples of isolated strong lensing ETGs in the Sloan Lens

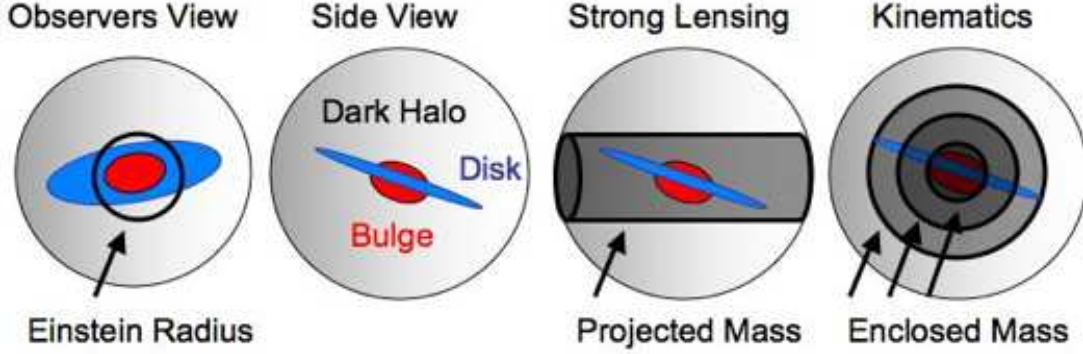


Figure 1.4: Strong lensing and dynamics provide complementary constraints on galaxy mass, since strong lensing probes the projected mass within a 2D aperture, whereas kinematics are sensitive to the 3D enclosed mass. This cartoon is reproduced from Dutton et al. (2011).

ACS Survey (SLACS; Bolton et al., 2008) and the Strong Lensing Legacy Survey (SL2S; Gavazzi et al., 2012) to robustly demonstrate that the slope of their *total* mass density is very close to isothermal (Treu & Koopmans, 2002; Koopmans & Treu, 2003; Koopmans et al., 2006, 2009; Barnabè et al., 2009; Auger et al., 2010a; Barnabè et al., 2011; Sonnenfeld et al., 2013). However, it is difficult to disentangle the dark and luminous mass components on these scales – especially using aperture mass measurements such as the central stellar velocity dispersion and Einstein radius – due to the absence of large-radius tracers. So far, strong constraints have only been obtained for one unusual isolated system – the compound lens J0946+1006 (Gavazzi et al., 2008; Sonnenfeld et al., 2012), in which the presence of lensed sources at different redshifts and hence with different Einstein radii allows a more radially extended probe of the projected mass – which appears to have a strongly contracted halo ($\gamma = 1.7 \pm 0.2$; Sonnenfeld et al., 2012). For ordinary (single source plane) isolated lenses, constraints from the hierarchical modelling of statistical samples are consistent with this result, though themselves weak (Sonnenfeld et al., 2015) unless some fixed IMF is assumed (Grillo, 2012).

On the scales of group- and cluster-lenses, on the other hand, the existence of mass probes at larger radii has allowed substantial progress to be made, and led to the emergence of a coherent – but not yet compelling – picture in which the importance of different baryonic processes depends on galaxy environment. On cluster scales, the simultaneous modelling of aperture mass measurements from stellar kinematics and strong and weak lensing found a sample of seven brightest cluster galaxies (BCGs) to generally have shallow

dark matter haloes compared to the NFW prediction (with a mean inner slope $\gamma = 0.5 \pm 0.1$; Newman et al., 2013); meanwhile, a similar study of 10 group-scale lenses found evidence for mildly contracted haloes (with inner slopes which are NFW-like within 2σ ; Newman et al., 2015). Taken together, these results tentatively suggest that the halo structure – and in turn, the relative importance of different baryonic processes – may depend on environment, such that galaxies in dense clusters are more strongly affected by heating due to satellite infall whereas group-scale lenses, in poorer environments, more strongly retain the signature of adiabatic contraction due to the initial inflow of gas. However, these two studies deal with small galaxy samples and use similar datasets and modelling assumptions, and further independent investigations are needed to verify the robustness of the results (see, for instance, Host & Hansen, 2011, for evidence that at least the *total* mass slope in X-ray clusters varies widely). Moreover, whilst the Sonnenfeld et al. (2012) result for the isolated double source plane lens seems to be consistent with this environmentally-dependent picture, better determinations of the mass structure of these lower-mass scales are essential if we are to test the hypothesis of an environmental dependence at these scales and understand the halo structure and the underlying physical processes which govern it.

1.4 This thesis: challenges in ETG evolution

In this thesis, I address three major questions regarding the evolution of ETGs.

1. What is the structure of the dark halo in ETGs?

Studies based on small samples of cluster-scale and group-scale ETGs suggest that halo structure is an environment-dependent property, but studies of lower-mass, isolated ETGs and studies using different techniques are needed to investigate these early results. In Chapters 2 and 3 I combine multiple dynamical tracers to disentangle the dark and light mass in the BCG M87 and so extend the results of Newman et al. (2013) to non-lenses; I then further demonstrate the robustness of this approach in Chapter 7. In Chapter 8 I extend constraints on the ETG halo structure to isolated lenses.

2. What is the nature of the IMF in ETGs?

The IMF in massive ETGs appears to be bottom-heavy relative to that of the Milky Way, but it is not clear whether this is globally true or confined to the central regions of these systems. Evidence for the existence of IMF gradients within individual ETGs so far has been restricted to stellar population analyses of ETG spectra, and has not been verified using dynamical or strong lensing probes. In Chapter 7, I present the first evidence for the existence of an IMF gradient in M87 using dynamics alone. In Chapter 8 I use strong lensing to constrain IMF gradients in 12 strong lenses.

3. What are the physical mechanisms driving ETG growth?

Evidence from the evolution of the ETG population implies that they have experienced significant size evolution since $z \sim 2$, but problems with our best current explanation, in terms of dry minor merging, suggest that other growth mechanisms may also be important. In Chapter 4, I investigate ETG size evolution using the FP, to consider possible differences in the evolutionary paths of ETGs since $z \sim 0.6$ in clusters as compared to in the field. In Chapters 5 and 6, I exploit the magnification of lensing to super-resolve compact ETGs at redshifts $z \sim 0.6$ in order to obtain new morphological evidence for their inside-out growth.

GALAXY STRUCTURE FROM MULTIPLE TRACERS: I. A CENSUS OF M87'S GLOBULAR CLUSTER POPULATIONS

Abstract

We present a new photometric catalogue of the rich globular cluster system around M87, the brightest cluster galaxy in Virgo. Using archival Next Generation Virgo cluster Survey (NGVS) images in the *ugriz* bands, observed with CFHT/MegaPrime, we perform a careful subtraction of the galaxy's halo light in order to detect objects at small galactocentric radii as well as in the wider field, and find 17620 globular cluster candidates over a radius range from 1.3 kpc to 445 kpc with magnitude $g < 24$. By inferring their colour, radial and magnitude distributions in a Bayesian framework, we find that they are well described as a mixture of two globular cluster populations and two distinct contaminant populations, but confirm earlier findings of radius-dependent colour gradients in both globular cluster populations. This is consistent with a picture in which the more enriched globular clusters reside deeper in the galaxy's potential well, indicating a role for dissipative collapse in the formation of both the red and the blue globular cluster populations.

2.1 Introduction

The globular cluster populations of a galaxy hold a wealth of information about the galaxy itself, both past and present. While the frequently observed bimodality in their colours (Zepf & Ashman, 1993; Ostrov et al., 1993; Gebhardt & Kissler-Patig, 1999) hints at a

non-trivial formation history, with at least two major formation periods and mechanisms at work, their extended spatial distributions make them good probes of the dark matter distribution at large radii, which can otherwise prove elusive. In this way, globular cluster dynamics can offer powerful insights into a galaxy's extended mass structure in a way that the much more centrally-concentrated starlight alone cannot.

However, any inference based on a subsample of a galaxy's globular clusters depends on a proper characterisation of the spatial profiles of the underlying population. For instance, the subsample of globular clusters in extragalactic systems for which we can obtain reliable spectroscopy tends to be subject to some non-trivial selection criteria, and this can lead to apparent spatial distributions which deviate dramatically from those of the parent populations. Dynamical models of the galaxy based on these spectroscopically-determined distributions rather than the true underlying ones can therefore result in very different conclusions. In the same way, as globular clusters generally populate their host galaxies with surface density profiles that fall off rapidly with galactocentric distance – often characterised, for instance, by Sérsic profiles – it is important that this characterisation of the underlying population is informed by globular clusters as close-in to the galaxy's centre as possible; this way, the innermost slope of the profile can be constrained much more precisely. These globular cluster populations (and planetary nebula populations, if present in significant numbers), with their distinct spatial and kinematic signatures, can then be used in dynamical models as independent tracers of the gravitational potential, providing larger-radius constraints on the galactic structure which are complementary to those from the starlight.

The massive elliptical M87, located at the centre of the Virgo cluster, is an ideal subject for dynamical globular cluster projects such as these. This (BCG) hosts the largest collection of globular clusters in the local Universe, with estimates as large as $N \sim 12,000$ (McLaughlin et al., 1994; Tamura et al., 2006a; Durrell et al., 2014). Its large globular cluster population was first recognised by Baum (1955), in a study which compared the globular cluster luminosities with those of M31 for use as a distance indicator. Since then, its globular clusters have been the target of a number of observational programmes, and recent work has conclusively shown that there are at least two distinct populations of globular clusters that are separated in both their colour and radial profiles, with the blue globular clusters extending significantly further than the red globular clusters (e.g. Harris, 2009). These multiple populations are important dynamical tracers of the potential well at different radii, and several groups have recently exploited the large catalogue of spectroscopic globular cluster velocities from Strader et al. (2011) to place constraints on the dark matter profile and stellar mass-to-light ratio of M87 (Agnello et al., 2014; Zhu et al., 2014).

However, this high-quality spectroscopy has yet to be coupled with a radially complete, publicly available photometric catalogue of the kind needed to properly characterise the underlying populations. Indeed, whilst a number of photometric catalogues exist, most have only partial radial coverage, and care must be taken to relate these catalogues when the observations are made using different filters and instruments. The HST catalogue of Peng et al. (2009), for instance, is extremely deep and complete out to a galactocentric radius of ~ 6.5 kpc in the ACS F606W and F814W filters, while NGVS provides archival catalogues covering a large extent of the Virgo cluster in the CFHT/MegaPrime *ugriz* bands, but with the regions immediately surrounding all bright objects (including M87) masked. The impetus for this study, then, is to bridge the gap and compile a comprehensive, extensive catalogue in a single filter system, by going back to the original NGVS images to carefully model and subtract the galaxy light and produce catalogues with more complete radial coverage. These can then be used to robustly characterise the globular cluster distributions, acting as a springboard for the dynamical study presented in Chapter 3.

The chapter is structured as follows: in Section 2.2, we introduce the dataset and explain our background subtraction. In Section 2.3, we present our photometry; Section 2.4 describes our globular cluster selection methods and Section 2.5 our inferred distributions. We discuss our findings in Section 2.6 and summarise in Section 2.7.

2.2 Data & reduction

We downloaded stacked NGVS images in the region around M87 from the Canadian Astronomy Data Centre (CADC). NGVS imaged a total of 104 deg^2 within the virial radii of Virgo’s A and B sub-clusters in the *ugriz* bands of the MegaPrime instrument on the Canada-France Hawaii Telescope (CFHT), over 200 nights between 2009-2014 (full details can be found in Ferrarese et al., 2012). The images available for download have a pixel scale of $0.187''/\text{pixel}$ and have been pre-processed using the (NGVS-tailored) Elixir-LSB pipeline – which includes bad pixel masking, bias and overscan subtraction, flat fielding and the removal of scattered light – and then photometrically and astrometrically calibrated and stacked. They are accompanied by exposure time and bad pixel maps and preliminary source catalogues obtained using SExtractor. We downloaded images for the four fields covering M87 and its surroundings: in the NGVS file-naming system (in which the first number refers to the RA offset from the survey centre and the second to the Dec equivalent), these are the +0+0, +0+1, -1+0 and -1+1 tilings. As noted in the Introduction, the part of the image immediately surrounding M87 is extremely bright due to M87’s stellar profile, making source detection and measurement incomplete and unreliable out to a radius of 6

kpc, and this region has also been removed from the NGVS catalogues.

Careful modelling and subtraction of the stellar light distribution allowed us to significantly improve the situation. We first performed an object detection step to avoid removing light from globular cluster candidates and thereby underestimating their flux. This was accomplished by first subtracting a heavily median-filtered version of the image from the original, then applying a clipped mean/variance filter to flag all pixels with flux 2σ above the local noise level, and finally applying a series of erosion and dilation filters to the flag image to remove noise and expand the mask around real objects. This mask was then applied to the original image, which allowed us to fit and subtract a spline model of the light profile in the radial direction. This removed a significant component of M87’s stellar light, as can be seen by comparing the first two panels of Figure 2.1, but as the distribution of the light is not purely radial, this left the image with the X-shaped pattern that is prominent in the central panel. To remove this effectively, we Fourier-filtered the high frequencies from the intermediate image to construct a smooth background, which could again be subtracted. The difficulty here is that Fourier filtering requires a real-space image with no gaps, i.e. no masked sources, so we iterated this procedure a number of times, each time updating the source mask and using the current best smooth background image to fill in the masked pixels and so converge to an acceptable approximation of the background.

As can be seen in the right-hand panel of Figure 2.1, the result of this step was a much cleaner image down to ~ 1.1 kpc, with a clear jet shooting off to the right and a large number of globular clusters revealed close to the galaxy’s centre. In all except the u -band (where the galaxy is not very bright), the very central region proved impossible to model due to saturation, and, having made this region as compact as possible, we masked it in the subsequent analysis.

2.3 Photometry

We used SExtractor and PSFex (Bertin, 1996) to detect and measure sources. For these routines to work in conjunction, SExtractor must be run first to provide a set of small images (‘vignettes’) for each detection. These vignettes are then used by PSFex to construct a model of the point-spread function (PSF) which can vary across the field of view, and which is fed into a final run of SExtractor to measure PSF-fitted magnitudes. This three-step process is preferable to running SExtractor alone when the field is crowded, as it enables a greater degree of deblending than could otherwise be achieved. SExtractor can also use a weight map alongside the detection and measurement images to deal with variations in noise, and here we found that the use of the NGVS exposure time maps was

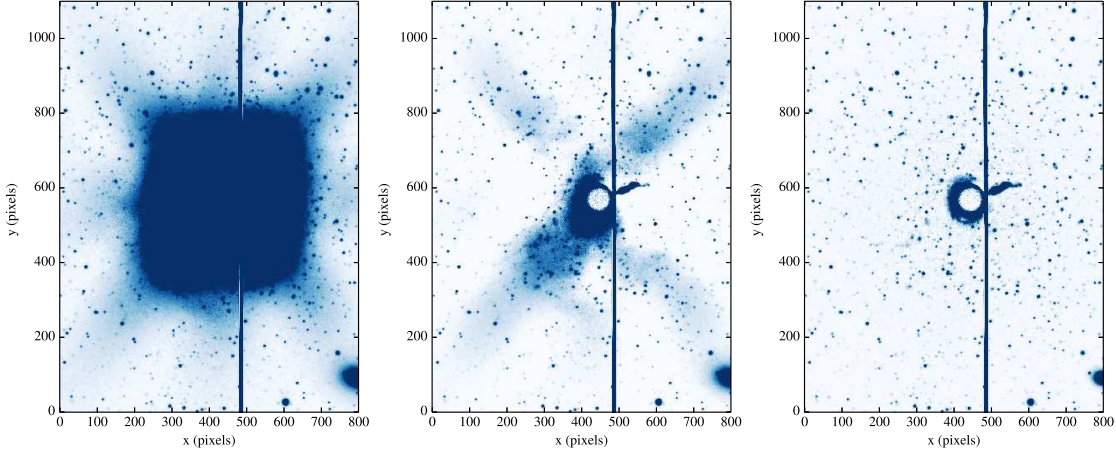


Figure 2.1: Step-by-step light subtraction in the r -band. Left: In the original NGVS image, the globular clusters close to M87 are totally dominated by the stellar light, making photometry unfeasible. Centre: Subtracting a spline model of the light in the radial direction helped to reveal the central region, but the lack of azimuthal symmetry in the original image resulted in a residual X-shaped pattern. Right: After the subtraction of a smooth Fourier-filtered background, the inner region of M87 is much cleaner. The circle in the very centre represents the saturated pixels, and is surrounded by a small bright ring which we were unable to totally eliminate, but whose size we substantially reduced; the saturated columns to the right of M87 are bleeding down from a star at higher declination. All three images are shown on the same scale.

not sufficient to handle the increased shot noise in the innermost regions of the galaxy as compared to the wider field. To account for this properly, we combined the exposure time maps with a second set of maps quantifying the relative pixel-to-pixel variance. We then ran the SExtractor-PSFex-SExtractor routine for each of the five bands and each of the four fields, using the g -band as the detection image as this had the most independent detections and so allowed us to obtain as complete a catalogue as possible. We added the uncertainties on the magnitudes output by SExtractor in quadrature with a systematic uncertainty of 0.02 mag to account for uncertainty in the photometric zeropoints. The use of a detection image made the merging of the different bands trivial; we also merged the catalogues from the different fields, removing duplicates in the overlapping regions by discounting any objects centred within 1.5 pixels of one another.

To remove regions of the image with unreliable photometry – primarily, those close to the bleed trails of very bright stars and their pupil ghosts – we made use of some of the other parameters output by SExtractor, identifying all objects fulfilling the following criteria in the g -band, and applying a dilation filter to a map of their positions to create a

mask:

1. FLUX_RADIUS >10 pixels: SExtractor measures a half-light radius based on the flux inside a (user-defined) circular aperture, which can be used to identify objects likely to be extremely bright foreground stars
2. MAG_AUTO <16 mag: this magnitude measurement, made using flexible apertures centred on each detection prior to PSF-fitting, is preferable to the PSF-fitted magnitude for identifying bright extended objects, whose fluxes are likely to be underpredicted by the PSF model.

Based on this masking, we calculated the fractional effective unmasked area $\frac{A_{eff}}{A}(R)$ of our field of view as a function of galactocentric radius and fitted it with a spline; this is important for relating measured number counts to physical surface densities. Explicitly, the intrinsic and observed number counts are simply related by

$$(2.1) \quad \begin{aligned} n_{int}(R) &= \frac{n_{obs}(R)}{A_{eff}/A(R)} \\ N_{int}(R) &= \frac{N_{obs}(R)}{A_{eff}/A(R)} \end{aligned}$$

for surface number density n and total number N .

To test our photometry, we cross-matched our objects with those in the NGVS and HST catalogues. When comparing with NGVS, a small amount of scatter is to be expected due to differences in our detection methods – the NGVS catalogues were generated using the exposure-time weight maps only, and without modelling the PSF – but the overall scatter is just ~ 0.017 mag down to 24th magnitude. We performed a similar comparison with the HST/ACS B,V,I photometry in the catalogue of Peng et al. (2009), this time selecting globular cluster candidates as explained in Section 2.4 and calculating synthetic photometry based on a mixture model of four single stellar populations (SSPs), at different ages and metallicities to reflect the bimodal nature of the globular cluster population. The scatter in this case is still consistent with experimental uncertainty at ~ 0.03 mags, especially given the simplicity of our four-component SED model. We also used the HST image to measure the completeness of our catalogue, given the depth and extremely high resolution of the former. Using our best-fit synthetic photometry, we confirmed that our catalogue remains complete down to 24th magnitude, with no dependence on radius outside the masked central ~ 1 kpc region.

As a further check that the background subtraction around M87 had not biased our photometry, we ran a series of simulations, synthesising stellar objects with Gaussian

profiles and known magnitudes and inserting them into the original pre-subtraction image, then running the subtraction procedure and comparing the magnitudes output from SExtractor with their known magnitudes. We did this for a total of 100 sources, implanting groups of 10 to avoid dramatically changing the density, and found that SExtractor managed to consistently reproduce the magnitudes to within a maximum difference of 0.04 mags. We also ran a similar experiment to test SExtractor’s detection efficiency in the central regions around M87, inserting synthetic objects exclusively in this area; SExtractor was able to detect and recover unbiased photometry for all sources.

2.4 Globular Cluster Candidates

The line of sight to M87 is heavily contaminated by stars in the Milky Way halo (and, to a lesser extent, the thick disk) and the Sagittarius stream, and is seen against a backdrop of interloping galaxies. We identified globular cluster candidates according to the sizes, colours and magnitudes of sources via the following steps:

1. *Selecting point sources:* Plotting magnitude against SExtractor-measured sizes, a clear horizontal distribution of unresolved objects emerges. Treating each field separately to allow for variations in the PSF, we drew stellar selection boxes around these branches in the g -band images, with a consistent faint-end magnitude cut of $g < 24$ mags to limit contamination. The upper limit on the radius was determined by considering the sizes of the objects classified as globular clusters in Strader et al. (2011)’s kinematic sample and ensuring that all of these, except for a few extreme outliers, survived the cull. An example of this selection is shown in Figure 2.2 for the +0+0 field. Our selection is deliberately conservative, since non-globular cluster contamination is dealt with in the subsequent analysis.
2. *Colour selection:* Because of the bimodal nature of M87’s globular clusters, they lie on two distinct branches in colour-magnitude space, as can be seen in Figure 2.3. The diagram has an overdensity at $g-i \sim 0.4$ due to main-sequence turn-off (MSTO) stars in the intervening Milky Way halo and, on the red end at $g-i \sim 2-2.5$, a contribution from nearby, low-luminosity disk stars, while unresolved background galaxies span colour space at the faintest magnitudes. This leaves two distinct peaks at $g-i \sim 0.7$ and $g-i \sim 1.0$ which we identify as the blue and red globular cluster populations, though we note that the peak colour of the interloping Sagittarius stream also lies close to the blue globular cluster population (although these stars are typically

brighter than globular clusters). We therefore isolate the globular cluster branches by imposing the cuts:

$$\begin{aligned} 0.2 < g - r < 1.0 \\ 0.5 < g - i < 1.45 \\ i > 18.0. \end{aligned}$$

These final cuts leave a catalogue of 17620 globular cluster candidates spanning radii from 1.3 kpc to 445 kpc, though the azimuthal coverage is only complete out to 240 kpc. We therefore define a second catalogue of 10784 objects which extends out to this radius. This is the catalogue which we use in the analysis that follows. A sample of the full catalogue is presented in Table 2.1, and the full version is available online¹. The photometry provided in the catalogue has not been corrected for dust extinction, though we do correct for this in our analysis, using the dust maps of Schlafly & Finkbeiner (2011).

2.5 Globular Cluster Populations

The globular clusters in M87 comprise multiple separate populations with distinct colours, spatial distributions, globular cluster luminosity functions (GCLFs) and formation histories, the standard scenario being a two-component model with a redder, more compact (and usually referred to as metal-rich) population existing alongside a bluer, more extended one (e.g. Tamura et al., 2006b). These differences act as filters to help in picking out the globular cluster populations from the interloping objects. We therefore model the catalogue as being composed of four distinct components, comprising the red and the blue globular cluster populations and two contaminant populations. We allow for one population of interloping stars from the Milky Way disk and halo and the Sagittarius stream, whose colour distributions we can model in detail, and a second interloping population of uniform colour, which could include distant background galaxies and stars from other sources (e.g. the Virgo overdensity). Within this paradigm, each globular cluster population follows a Sérsic profile in radius, a Gaussian distribution in $g - r$ and $g - i$ colours and a Gaussian GCLF in the g -band. This form for the GCLF was chosen to facilitate comparison with previous studies (e.g. Hanes, 1977; Tamura et al., 2006a; Harris et al., 2014). Likewise, we chose to use Sérsic profiles in radius following previous authors (Strader et al., 2011; Agnello et al., 2014). We also allow for radial gradients within the colour distributions of each globular cluster population, using the functional form

$$(2.2) \quad \mu' = \mu - \nu \log(R/R_{eff})$$

¹<http://adsabs.harvard.edu/abs/2015yCat..74550820O>

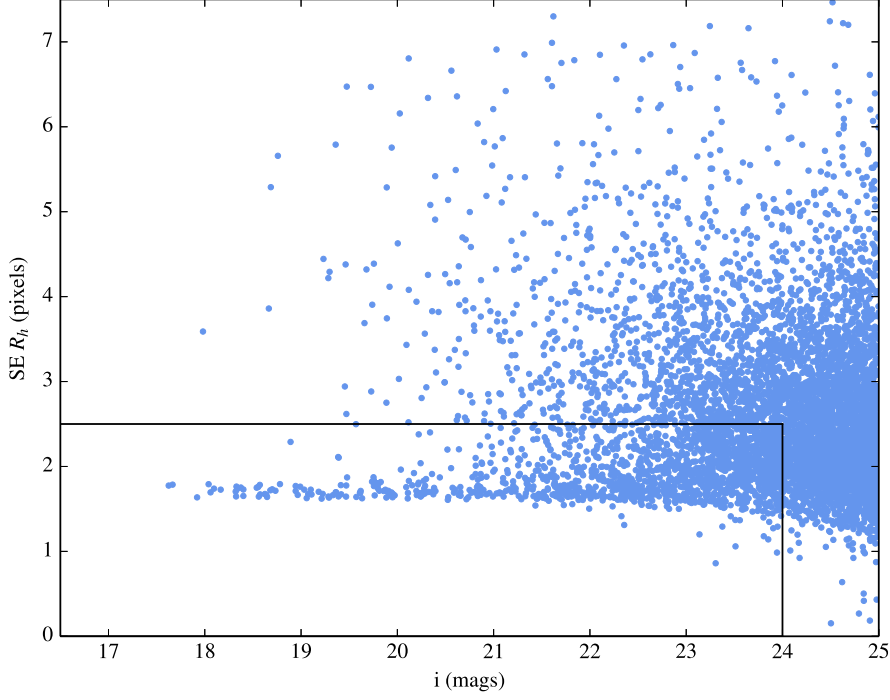


Figure 2.2: Point source selection using SExtractor’s effective radius parameter: SExtractor measures a half-light radius for each detection based on its fixed-aperture magnitudes, which, in the case of point sources, can be used as a proxy for the PSF. Point sources therefore lie on a distinct branch of small radius with low scatter. For each field, all objects outside the stellar selection box (defined using the g -band image) were rejected from the sample. At the faint end of the box, we chose our magnitude cut-off such that our catalogue depth would be comparable to the turnover magnitude of the globular cluster populations and would include the main population of Strader et al. (2011)’s kinematic sample. Clearly, this choice of magnitude cut-off is a compromise between the detection of fainter globular clusters and the amount of contamination in the catalogue. We also note that this magnitude cut-off causes the completeness to be a function of colour, as can be seen in Figure 2.3.

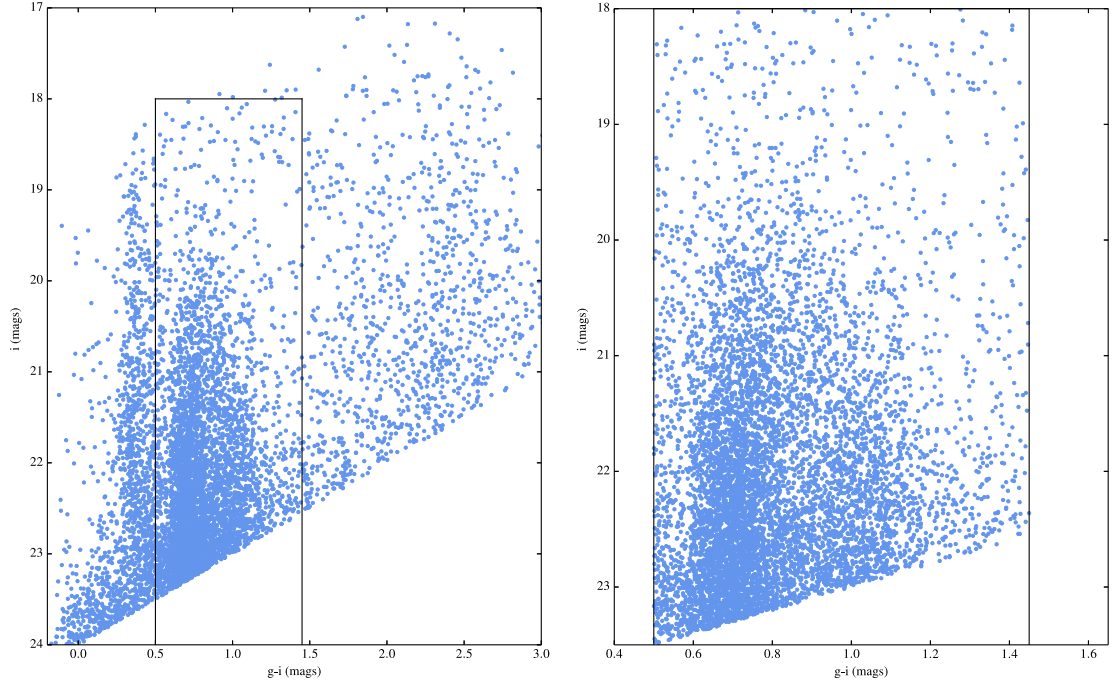


Figure 2.3: Colour selection: M87's globular clusters are known to be bimodal in colour space, as can be seen particularly clearly in the $g-i$ CMD. Left: The distribution of all point sources, with our selection limits overplotted. The bluest population at $g-i \sim 0.4$ is made up of MSTO stars in the Milky Way halo, while the extremely red component is due to low-luminosity disk stars. Bracketed by these is the double-peaked globular cluster distribution, although note that stars from the Sagittarius stream also have very similar colours to the blue globular cluster population. Right: The $g-i$ CMD after all colour cuts have been applied. Note that the completeness is a function of colour due to the g -band magnitude cut.

RA (deg)	Dec (deg)	<i>u</i> (mag)	<i>g</i> (mag)	<i>r</i> (mag)	<i>i</i> (mag)	<i>z</i> (mag)	p(red)	p(blue)	p(MW)	p(int)	flag
187.8039	11.4863	24.59±0.09	23.93±0.03	23.62±0.05	23.25±0.05	23.25±0.14	< 0.01	0.44±0.02	< 0.01	0.56±0.02	0
188.0390	11.4866	25.51±0.21	23.99±0.03	23.48±0.04	23.18±0.05	23.26±0.14	0.22±0.03	0.06±0.03	< 0.01	0.71±0.02	0
188.0177	11.4882	24.21±0.07	22.30±0.02	21.56±0.02	21.20±0.02	20.91±0.03	< 0.01	< 0.01	< 0.01	> 0.99	0
187.6281	11.4881	24.45±0.08	23.80±0.03	23.11±0.03	22.97±0.04	22.75±0.09	0.03±0.02	< 0.01	< 0.01	0.97±0.02	0
187.7569	11.4880	24.49±0.09	23.44±0.03	22.65±0.03	22.47±0.03	21.81±0.04	< 0.01	< 0.01	< 0.01	> 0.99	0
188.3855	11.4890	25.11±0.14	22.99±0.02	22.21±0.02	21.77±0.02	21.52±0.04	< 0.01	< 0.01	< 0.01	> 0.99	0
188.1248	11.4884	22.81±0.03	21.75±0.02	21.30±0.02	21.10±0.02	20.94±0.03	< 0.01	0.94±0.01	< 0.01	0.06±0.01	0
187.8688	11.4890	24.28±0.07	23.27±0.02	22.90±0.03	22.71±0.03	22.71±0.09	< 0.01	0.78±0.03	< 0.01	0.22±0.03	0
188.3448	11.4924	23.22±0.03	21.27±0.02	20.37±0.02	20.02±0.02	19.75±0.02	< 0.01	< 0.01	0.61±0.01	0.39±0.01	0
188.4594	11.4954	24.65±0.07	23.84±0.03	23.44±0.04	23.33±0.05	23.01±0.12	< 0.01	0.24±0.01	0.65±0.01	0.11±0.01	0
188.2633	11.4950	22.59±0.02	21.56±0.02	21.07±0.02	20.84±0.02	20.60±0.02	0.11± 0.01	0.27±0.01	< 0.01	0.62±0.02	0
187.8758	11.4946	20.89±0.02	19.77±0.02	19.31±0.02	18.95±0.02	18.82±0.02	0.07±0.02	<0.01	< 0.01	0.93±0.02	0
187.6205	11.4951	20.48±0.02	19.40±0.02	19.05±0.02	18.77±0.02	18.59±0.02	< 0.01	0.67±0.03	< 0.01	0.31±0.03	0
187.6787	11.4968	23.98±0.04	22.94±0.02	22.56±0.03	22.21±0.03	21.94±0.04	0.01±0.01	0.54±0.02	< 0.01	0.46±0.02	0
187.8992	11.4983	24.68±0.07	23.59±0.03	23.29±0.03	22.93±0.04	22.80±0.08	< 0.01	0.21±0.03	< 0.01	0.79±0.03	0
187.9040	11.4985	24.68±0.07	23.72±0.03	23.47±0.04	23.17±0.04	22.96±0.09	< 0.01	0.01±0.01	< 0.01	0.99±0.01	0
188.3613	11.5001	24.24±0.06	23.30±0.02	23.01±0.03	22.68±0.03	22.50±0.06	< 0.01	0.09±0.02	< 0.01	0.91±0.02	0
187.9755	11.5010	24.03±0.04	22.97±0.02	22.58±0.03	22.16±0.03	21.85±0.04	0.01±0.01	0.01±0.02	< 0.01	0.98±0.02	0
187.9066	11.5004	24.84±0.08	23.65±0.03	23.22±0.03	22.92±0.04	22.60±0.07	0.01±0.01	0.76±0.01	< 0.01	0.23±0.01	0
187.8627	11.5008	22.28±0.02	21.26±0.02	20.87±0.02	20.49±0.02	20.38±0.02	0.01±0.01	0.07±0.03	< 0.01	0.93±0.03	0
188.4949	11.5020	22.86±0.02	22.00±0.02	21.69±0.02	21.43±0.02	21.33±0.03	< 0.01	0.02±0.01	< 0.01	0.98±0.01	0
188.3620	11.5033	24.91±0.08	23.71±0.03	23.34±0.04	22.92±0.04	22.72±0.07	0.01±0.01	0.08±0.02	< 0.01	0.92±0.02	0

Table 2.1: Catalogue data. Magnitudes are PSF-fitted using SExtractor and PSFex, and are measured in the CFHT/MegaPrime filter system without correcting for dust extinction. Columns 1 and 2 list the RA and Dec of the sources; columns 3 - 7 present the *ugriz* magnitudes with associated uncertainties; columns 8 - 11 give the the probabilities of belonging to each of the four populations with associated uncertainties, and column 12 gives the object's velocity if it is included (and classified as a globular cluster) in the kinematic catalogue of Strader et al. (2011) and is set to zero otherwise. The full version of this table is available in machine-readable form at <http://adsabs.harvard.edu/abs/2015yCat..745508200>.

where μ and ν define a log-linear relation between radius and the peak μ' for each of the $g-i$ and $g-r$ Gaussian colour distributions, with R_{eff} set equal to 16 kpc, the effective radius of the starlight as reported in Kormendy et al. (2009).

The uniform-colour interloping population is also uniform in space, and has a luminosity function (LF) based on the form of the catalogue's LF at large radii, where the globular cluster profiles are assumed to have largely died away. To check this was a reasonable assumption, we binned the full catalogue radially and compared how its LF changed as a function of radius in the outermost bins. As Figure 2.4 shows, there is negligible variation at large radii. Finally, the Milky Way interlopers have a uniform spatial distribution – as we are only sampling small fractions of the Sagittarius/Milky Way systems in our field of view – and a colour distribution based on a combination of synthetic survey data for the Milky Way (using the code Galaxia, Sharma et al., 2011) and for Sagittarius using models for the star formation history from observations in SDSS Stripe 82 (de Boer et al., 2015). Whilst the radial and colour distributions are the strongest diagnostics here, the relatively faint magnitude limit of our catalogue makes the LF a useful additional tool for deselecting contaminants, whose densities are expected to increase rapidly at the faint end.

We infer the parameters of our model in a Bayesian way such that, given data \vec{D}_k and model parameters \vec{M} , the posterior distribution for the model is given by the product

$$(2.3) \quad P(\vec{M}|\vec{D}_k) \propto P(\vec{D}_k|\vec{M}) P(\vec{M})$$

where the first term on the right represents the likelihood of the data given the model and the second term is the prior, which we assign to the model based on our existing knowledge. In our model, the data \vec{D}_k comprises a vector containing the galactocentric radius, the g -band magnitude and the $g-i$, $g-r$ colours,

$$(2.4) \quad \vec{D}_k = \begin{pmatrix} R_k \\ g_k \\ (g-i)_k \\ (g-r)_k \end{pmatrix}$$

and the model parameters \vec{M} are listed in Table 2.2. Assuming flat priors on all parameters, we can absorb these into our proportionality constant to write the posterior as

$$(2.5) \quad P(\vec{M}|\vec{D}_k) \propto P(\vec{D}_k|\vec{M}),$$

and, as each catalogue entry constitutes an independent observation, we can write the joint posterior distribution for the k objects in the catalogue as the product

$$(2.6) \quad P(\vec{M}|\vec{D}) = \prod_k P(\vec{M}|\vec{D}_k)$$

where \vec{D} now represents the full dataset.

With this formalism established, we now turn to the detailed form of the likelihood function $P(\vec{D}_k|\vec{M})$ of observing a single object given a set of model parameters. As explained at the beginning of this Section, our model comprises four distinct components; as such, the likelihood function for the k^{th} object is simply their weighted sum,

$$(2.7) \quad P(\vec{D}_k|\vec{M}) = (f_r P_r + f_b P_b + f_s P_s + f_i P_i) \frac{A_{eff}}{A},$$

where the subscripts r , b , s and i correspond to the red globular cluster, blue globular cluster, interloping stellar and unclassified interloper populations respectively, and the f coefficients represent their relative fractions, normalised such that

$$(2.8) \quad f_r + f_b + f_s + f_i = 1$$

within the area covered by our data, i.e. out to 240 kpc. The factor $\frac{A_{eff}}{A}$ is the fractional effective area of our detection image at that object's galactocentric radius, accounting for the masking of bad pixels and bright objects, as explained in Section 2.3. In the log-likelihood calculation, this factor just gives a constant additive term, but it is important for comparing our radial profile with other studies, as it allows us to simply rescale our number count to account for the masked regions.

The likelihoods P_r , P_b , P_s , and P_i for the four populations are then the functions of R , $g-i$, $g-r$, and g described at the beginning of this Section, and can be summarised as follows:

$$(2.9) \quad \begin{aligned} P_r &= \Sigma(R|R_{e,r}, n_r) N(\vec{m}_r, R|\vec{\mu}_r, \sigma_r^2) \\ P_b &= \Sigma(R|R_{e,b}, n_b) N(\vec{m}_b, R|\vec{\mu}_b, \sigma_b^2) \\ P_s &= U(R) S(gi, gr, g) \\ P_i &= U(R, gi, gr) L(g) \end{aligned}$$

for Gaussian distributions N , uniform distributions U , Sérsic profiles Σ and the spline representations of the Milky Way stellar density distribution $S(gi, gr, g)$ and the uniform interloper luminosity function $L(g)$. Here, the red and blue globular cluster Gaussians are 4D, with, for instance, the centre of the red distribution given by $\vec{\mu}_r = (\mu_{gi,r}(R), \mu_{gr,r}(R), \mu_{g,r})$ with the radial dependence of the colours explicitly noted. Note that the GCLF Gaussian is explicitly truncated at $g = 24$ to account for our (imposed) selection function, which is assumed to be flat otherwise (note also that this accounts for the completeness being a function of colour, which, as explained in the captions of Figures 2.2 and 2.3, is a direct consequence of the magnitude cut).

The final model then has 23 free parameters, which we explore using the ensemble-sampling code EMCEE (Foreman-Mackey et al., 2013). Our inferred posteriors are shown

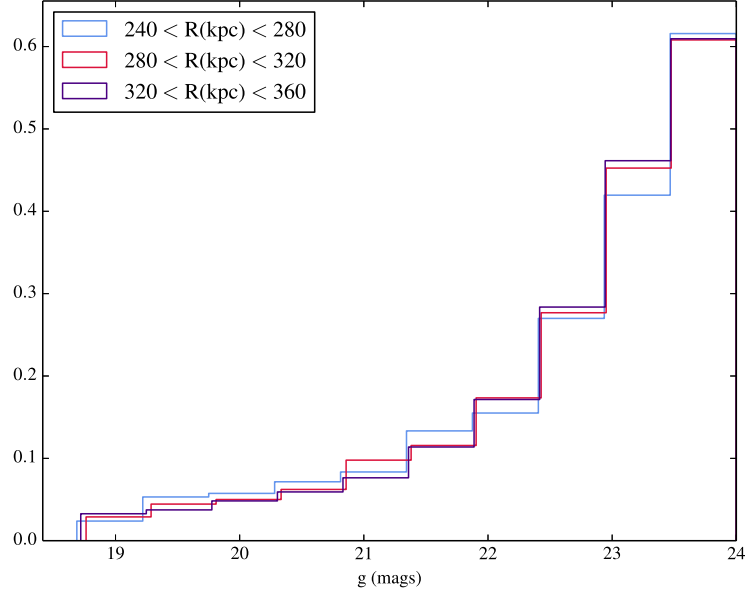


Figure 2.4: We based our model for the LF of the uniform-colour interloping populations on the LF of the full globular cluster candidate catalogue at large radii. The assumption is that the globular cluster profiles have largely died away by the time we reach these radii, meaning we can attribute the LF here entirely to the interlopers: otherwise, using this as our LF could artificially suppress the sizes of the globular cluster populations in our inference. These histograms show the LF in the outermost bins, and we see that the evolution is minimal.

in Figure 2.5, with their maximum-likelihood values and associated uncertainties listed in Table 2.2.

2.6 Discussion

Our inferred surface density profile is shown in Figure 2.5, and the parameters of the fit can be found in Table 2.2. Below, we discuss these results in more detail.

2.6.1 Characterising M87's globular cluster populations

In agreement with previous studies (e.g. Côté et al., 2001; Durrell et al., 2014), we find that the blue population within our catalogue is significantly more populous than the red, with the total globular cluster count made up of $\sim 1/3$ red globular clusters and $2/3$ blue

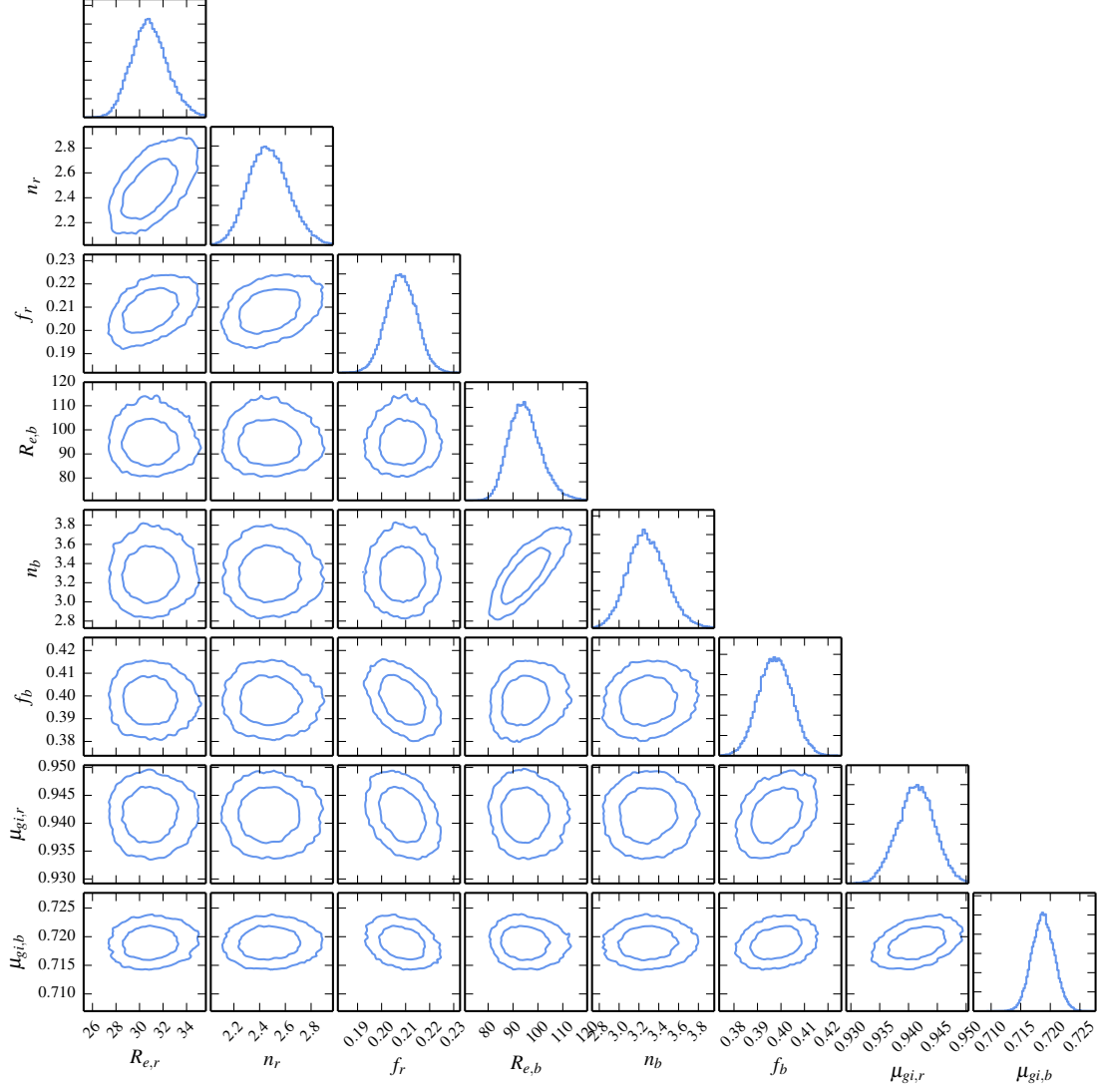


Figure 2.5: Inference on the model parameters. Our 23-parameter space includes Sérsic radial profiles and Gaussian colour distributions for the red and blue globular cluster populations, plus separate components of interloping stars and galaxies. Here we show the posterior marginalised over a number of parameters, in order to emphasise those dictating the globular cluster profiles, with contours representing the 68th and 95th percentiles.

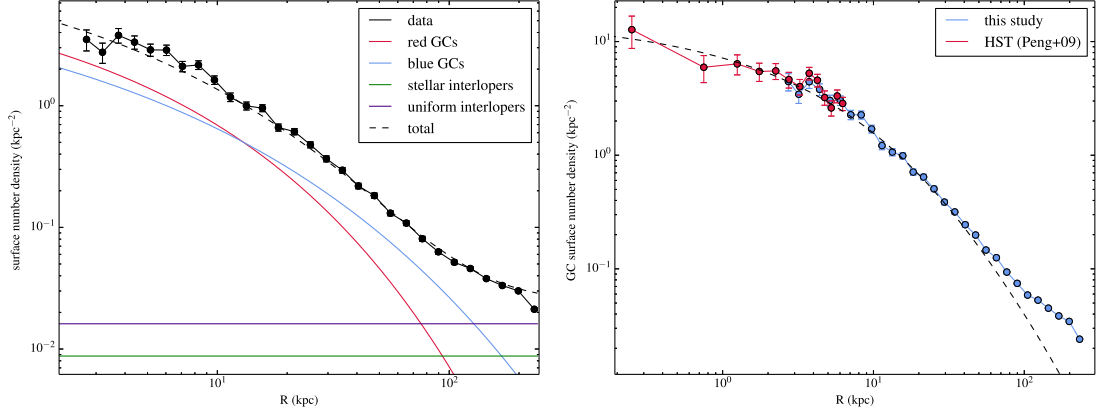


Figure 2.6: Left: Inferred surface density. The contributions from the red globular cluster, blue globular cluster and interloping components are plotted, together with the total inferred surface density, shown by the black dashed line. The data are overplotted in circles. Right: Comparison with the HST catalogue of Peng et al. (2009). These data, which extend to smaller radii than our catalogue, continue to follow our inferred globular cluster profile. Note that the surface density of objects in our catalogue deviates from the globular cluster profile at large radii because of the increasing relative importance of the interloper population; conversely, the dominance of the globular cluster populations at the smallest radii means our data can be well-approximated by the globular cluster curves alone in this region.

out to our cut-off radius of 240 kpc and down to 24th magnitude in the g -band. The blue population is also much more extended. However, given that the LF of the red population is roughly one magnitude fainter than that of the blue, we expect the red population to be larger overall: assuming our inferred distributions remain valid at large radii and faint magnitudes, we find that there should be ~ 6030 (mostly faint) red globular clusters and ~ 5300 blue globular clusters, giving 11330^{+1500}_{-300} globular clusters in total. This is consistent with most previous estimates (e.g. Tamura et al., 2006a), though it is smaller than the number count of $N = 14,520 \pm 1190$ reported by Durrell et al. (2014); however, the latter was a cluster-wide globular cluster survey and did not include the contaminant model for Sagittarius and Milky Way stars that we implement here, so it is possible that the blue globular cluster population derived in that study may include a contribution from Sagittarius stars, whose colour-magnitude distribution is very similar to that of the blue population. This possibility would also explain the fact that their number counts at large radii are slightly larger than what our Sérsic profiles predict. On the other hand, given that their data extend further out than the catalogue we use here, it is also possible that the blue population could shift to a different profile at these large radii. If this could be

		radial gradient model	fixed-Gaussian model
Red globular clusters	R_e (kpc)	30.6 ± 1.6	24.1 ± 0.9
	n	2.41 ± 0.16	2.012 ± 0.12
	μ_{gi} (mags)	0.944 ± 0.003	0.919 ± 0.003
	σ_{gi} (mags)	0.084 ± 0.002	0.100 ± 0.002
	μ_{gr} (mags)	0.597 ± 0.002	0.588 ± 0.002
	σ_{gr} (mags)	0.045 ± 0.002	0.055 ± 0.002
	f_r	0.204 ± 0.009	0.217 ± 0.009
	μ_g (mags)	24.60 ± 0.38	24.68 ± 0.29
	σ_g (mags)	1.65 ± 0.13	1.72 ± 0.10
	ν_{gi} (mags)	0.14 ± 0.01	—
	ν_{gr} (mags)	0.038 ± 0.003	—
Blue globular clusters	R_e (kpc)	95.0 ± 6.8	109.0 ± 8.5
	n	3.23 ± 0.2	2.74 ± 0.19
	μ_{gi} (mags)	0.719 ± 0.002	0.677 ± 0.001
	σ_{gi} (mags)	0.046 ± 0.001	0.051 ± 0.001
	μ_{gr} (mags)	0.459 ± 0.002	0.447 ± 0.002
	σ_{gr} (mags)	0.034 ± 0.002	0.036 ± 0.002
	f_b	0.398 ± 0.007	0.410 ± 0.007
	μ_g (mags)	23.65 ± 0.10	23.77 ± 0.14
	σ_g (mags)	1.37 ± 0.05	1.42 ± 0.06
	ν_{gi} (mags)	0.080 ± 0.003	—
	ν_{gr} (mags)	0.023 ± 0.003	—
Interlopers	f_s	0.134 ± 0.005	0.134 ± 0.005
	f_i	0.264 ± 0.006	0.239 ± 0.006

Table 2.2: Inferred parameters, for both fixed colour Gaussians and colour distributions whose mean decreases as a function of radius. Listed are the maximum-likelihood values from the pdf, along with uncertainties given by their 16th and 84th percentiles. All parameters were assigned uniform priors.

demonstrated, this would be an interesting result which could provide strong evidence for globular cluster accretion. However, whilst Durrell et al. (2014) note that the mean slope of the blue globular cluster surface density appears to become shallower at $R \sim 250$ kpc, they also stress that the uncertainties at these radii are too large to facilitate any strong conclusion.

Previous studies of globular cluster systems in elliptical galaxies have shown that the red globular cluster distribution tends to follow the starlight (e.g. Geisler et al., 1996), and this has been verified for M87 through studies such as that of Durrell et al. (2014). We test this for our model by comparing our Sérsic profile for the red globular clusters with a fit to the surface brightness distribution of the starlight given in Kormendy et al.

(2009), and find that the two profiles do indeed match closely at small and intermediate radii (< 100 kpc), but that the red globular clusters die away more rapidly than the stars at larger radii. As M87 is known to have an extended stellar envelope, possibly built up through merger events (e.g. Hausman & Ostriker, 1978), and virtually all globular cluster formation scenarios have the red population forming *in situ* (e.g. Ashman & Zepf, 1992; Côté et al., 1998), it is perhaps not surprising that this globular cluster population does not trace the starlight at these larger radii. We also note that, in spite of the increased depth of our catalogue, our average contaminant surface density of 0.58 arcmin^{-2} is similar to that found in e.g. Strader et al. (2011) and Tamura et al. (2006a), and that this increased depth is valuable in terms of globular cluster numbers; according to our model, we expect the globular cluster populations together to constitute $\sim 58\%$ of the objects with $g \sim 24$ out to 240 kpc. We compare the surface density of globular clusters as seen in the HST data of Peng et al. (2009) with our Sérsic profiles – shown in the right panel of Figure 2.5 – and find that our model is also a good fit at the innermost radii, where our catalogue does not reach. This is encouraging, and implies that any tidal disruption that may have occurred cannot have been effective in removing the central globular cluster population.

We find luminosity functions for the two globular cluster components whose peaks straddle the limiting magnitude of the catalogue, with $\mu_{g,r} = 24.60 \pm 0.38$ and $\mu_{g,b} = 23.65 \pm 0.10$ magnitudes, which are consistent with previous results. Tamura et al. (2006a) modelled the V -band distribution of the total globular cluster population using data extending out to ~ 0.5 Mpc and found a turnover magnitude of $V = 23.62 \pm 0.06$ and $\sigma_V = 1.40 \pm 0.04$, while McLaughlin et al. (1994) found a peak at $V = 24.2$ mags with a similar width of $\sigma_V = 1.73$. The GCLF is also consistent with that inferred by Peng et al. (2009) for globular clusters at smaller radii: they find a turnover magnitude for the total population $I = 22.53 \pm 0.05$, and when we cross-correlate the catalogues to calculate a colour correction, this translates to a g -band magnitude $g \sim 23.75$. Adopting the Tamura et al. (2006a) measurement of the total V -band galaxy luminosity of $M_V = -22.46$, we calculate the specific frequency S_N of M87's globular cluster system as 11.76 ± 2.1 , where we have included an uncertainty of 0.1 mags on the absolute magnitude M_V . As noted in other studies (e.g. McLaughlin et al., 1994), this is extremely high – the average S_N of other ETGs in Virgo is ~ 5 (Harris, 1991) and the specific frequency of the Virgo cluster as a whole is ~ 3 (Durrell et al., 2014) – and could be related M87's location at the bottom of the cluster's potential well, which may have led it to undergo an unusually large number of globular cluster-triggering mergers (e.g., Ashman & Zepf, 1992) or to have accreted many globular clusters from smaller satellite galaxies (e.g., Forte et al., 1982). Support for these environment-focussed interpretations comes from the fact that a number of other BCGs

have been found to have similarly high specific frequencies (e.g., McLaughlin et al., 1994).

We can also use the blue globular cluster population as an independent distance indicator by comparing the peak brightness of the GCLF with that of other metal-poor globular cluster systems. We take the turnover magnitude of these to be $M_V = -7.66 \pm 0.09$ (Di Criscienzo et al., 2006) and use the Peng et al. (2009) HST catalogue to determine an empirical $g - V$ correction of $g - V = 0.323 \pm 0.049$. This gives a distance modulus $DM = 30.99 \pm 0.14$, or $D_L = 15.79 \pm 1.04$ Mpc, in very good agreement with other measurements of M87's distance modulus (e.g., Ferrarese et al., 2000a; Mei et al., 2007).

M87's globular cluster population as seen in CFHT has previously been modelled by Strader et al. (2011), where the full photometric sample from CFHT was used to characterise the underlying distribution of a smaller kinematic subsample. This subsample was later used by Agnello et al. (2014) in a series of dynamical models, where they determined the apparent spatial distribution of the *kinematic* sample. The very significant difference between our Sérsic profiles and those in Agnello et al. (2014) – compare their $R_{e,r} = 6$ kpc and $R_{e,b} = 190$ kpc with our results in Table 2.2 – can be understood by the fact that the spectroscopic subsample incorporates some non-trivial selection function, meaning the radial distributions are indeed different. Our Sérsic profiles are similar to those inferred in Strader et al. (2011), though we find a Sérsic index for the red population which is significantly smaller than the $n_r = 5.33$ fitted in that work. It is important to note that our globular cluster populations differ from theirs slightly in that ours extend to a different magnitude limit (Strader et al., 2011, applied a magnitude limit of $i < 23$ mags in the SDSS i-band, which has only a small offset from the MegaCam i-band magnitude) and smaller galactocentric radii (in Strader et al., 2011, they fitted the data in radial bins starting at ~ 5 kpc); this could contribute in part to the difference, which in our models is driven by the objects at small radii, where the red globular clusters are most populous. Further, as noted in that paper, it is hard to assess the uncertainties on the Sérsic parameters due to their strong correlations.

2.6.2 Importance of correctly modelling the interlopers

An additional feature of our study is the more accurate and physically-motivated model that we use for the interloper colour distribution. Whilst the standard way to treat background contamination in these systems in the past has been to distribute them uniformly in both space and colour, the latter assumption is not justified for the case of intervening Milky Way and stream stars, which we should expect to have colours strongly clustered around the MSTO point and the red-giant branch (RGB) with a scatter determined by the spread

in stellar age. This gives a particularly strong bias in the case of Sagittarius, whose RGB lies virtually on top on the blue globular cluster distribution in colour space (as shown in Figures 2.7). In our model, then, what a uniform-colour contaminant model would assign to the blue population at large radius, we might be more likely to classify as an interloping star.

To test the sensitivity of the inference to different interloper distributions, we performed the inference with the following models:

1. a single interloper population, uniform in both space and colour
2. a single population of interloping (Milky Way and Sagittarius) stars
3. two populations, but with the contribution from Sagittarius excluded from the stellar model
4. a single population of Milky Way stars.

We found that case (i), assuming only uniformly distributed contaminants, required the scale radius and Sérsic index of the blue globular cluster population to be large: $R_{e,b} \sim 250$ kpc, $n \sim 3.5$. As explained above, this is due to the overlap of the blue globular clusters and the Sagittarius stars in the CMD. Case (ii), requiring all the interlopers to obey the Milky Way/Sagittarius model, converged on a solution in which Sagittarius dominated the catalogue to give $R_{e,b} \sim 30$ kpc, clearly highlighting the importance of including both sources of contamination. In case (iii), the lack of Sagittarius resulted in the stellar interloper distribution being a poor description of the data, such that the stellar fraction $f_s \rightarrow 0$ and the Sérsic profiles of case (i) were returned. Finally, case (iv) recovered profiles comparable to those obtained from the Sagittarius+Milky Way+uniform fits, with the absence of Sagittarius balanced, to some extent, by the absence of the uniform component. However, the fits in colour space were poor due to the mismatch of the Milky Way stars to the underlying contaminant distribution: in particular, the colour model had a bump at the red end due to low-luminosity disk stars, and this was obviously discrepant with the data. This, together with the significant f_s component that we infer, shows that the classification of objects as globular clusters versus contaminants is strongly dependent on our choice of contaminant model. Further, as the Sagittarius stellar population is generally bright, the peakiness of its distribution in colour space becomes more pronounced as the depth of the catalogue is reduced, making it even more important in shallower catalogues. Its exclusion could lead to dramatic overestimations of the extent of the blue globular cluster population, as well as how steeply it declines as a function of radius – which could explain, for instance, the larger total number count found in Durrell et al. (2014). This is something that must be modelled carefully in order to properly separate the populations.

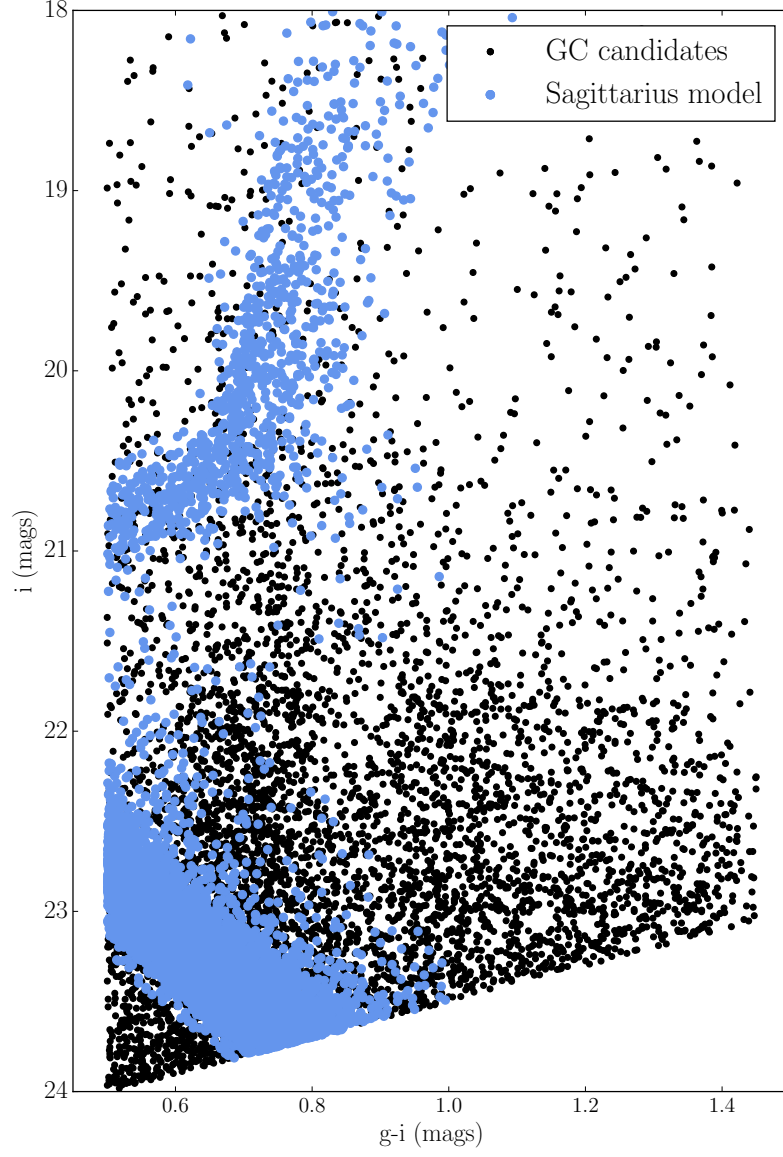


Figure 2.7: A comparison of the colour-magnitude distributions for our globular cluster candidates and our model of the Sagittarius stream shows the strong overlap between the latter and the blue globular cluster population. For this reason, a realistic model of the interloper colours is extremely important for our inferences on the distribution of the blue globular clusters.

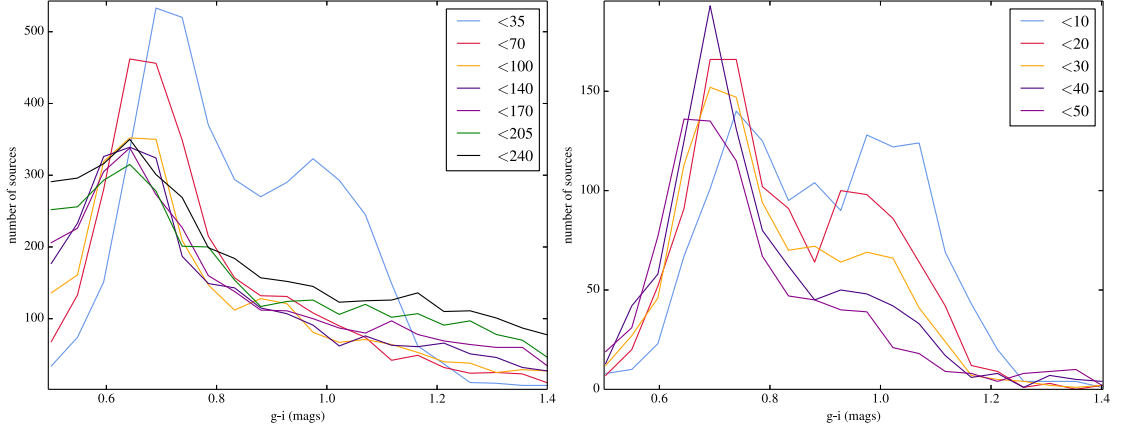


Figure 2.8: With increasing galactocentric radius, we see both the red and blue peaks migrating towards bluer colours. The histograms show the colour distributions of the binned data, with the legend denoting the maximum radius of each bin in kpc. In the right panel, we have restricted the dataset to $R < 50$ kpc to accentuate the red population.

2.6.3 The existence of colour gradients in the globular cluster populations

We now examine the cause of the colour structure within each globular cluster population as a function of radius, which was originally noticed in Strader et al. (2006) and recovered in Harris (2009) and Strader et al. (2011). In those works, the $g - i$ peak within each population was found to shift towards bluer colours as the radius was increased, though while Harris (2009) found this for both the red and the blue globular cluster populations, Strader et al. (2011) only found a significant trend in the red population. In our data, we find a gradient in the colours of both the red and the blue components as a function of galactocentric radius, and plot the colour histograms in radial bins in Figure 2.8 as a simple way of highlighting this.

In the modelling of Section 2.5, we interpreted these gradients as existing *within each population*, using Gaussian distributions whose peaks were allowed to vary as a function of radius in a log-linear fashion as in Equation 2.2. This is in line with Harris (2009), where the observed shift was used to argue for metallicity gradients within each population, which could in turn reflect the enrichment histories of the globular clusters. However, this is only one interpretation of the shift: Strader et al. (2011) used the same observation to motivate the existence of a third globular cluster population at intermediate colours, with the gradient arising as a result of the different populations merging into one another. Clearly, each of these scenarios has very different implications in terms of the properties

and formation histories of the globular clusters themselves. To distinguish between them, we also performed the inference without allowing for radial gradients, instead forcing the colour distributions to follow Gaussians with fixed peaks (we call this the ‘fixed-Gaussian’ model). We can then compare the inferred parameters for this model to the data, noting that the presence of a significant unmodelled third population would create substantial residuals between the data and the best-fit model in both the colour distribution and the radial profile, whereas colour gradients within the populations would largely affect the colour distribution only.

Figure 2.9 shows the colour distributions inferred for both the fixed-Gaussian and the radial gradient models. In the former, we see a discrepant bump at intermediate colours which the model cannot recreate. As explained above, this could be consistent with either colour gradients or a third population. The radial profile inferred for this model, on the other hand, is virtually unchanged from the previous ‘radial-gradient’ model, and does not show any evidence for systematic residuals. This suggests that the globular clusters are adequately described by a model with only two radial components, but not by two fixed-peak Gaussians in colour, a state of affairs which is more in line with the metallicity-gradient scenario than the three-population hypothesis. A further problem with the three-population scenario is that both our globular cluster populations exhibit colour gradients in the same direction (towards bluer colours). By the reasoning that leads to adding a third population, then, we should also be compelled to add a fourth, and the scenario quickly becomes more complicated.

As discussed in Harris (2009), the presence of metallicity gradients within the globular cluster populations has important implications for our understanding of the formation and enrichment histories of the globular clusters. That is, a population which becomes increasingly enriched at smaller radii, deeper in the gravitational potential well of the galaxy, is a signature of star formation via dissipative collapse, in contrast to that triggered by major mergers and accretion, where material tends to be more spatially mixed. That these gradients appear to exist in both globular cluster populations therefore indicates that both featured some component of dissipative collapse, and that M87’s two globular cluster populations cannot have formed via wholly exclusive channels (e.g. Côté et al., 1998).

2.7 Conclusions

We have performed a careful subtraction of M87’s stellar light using CFHT/MegaPrime *ugriz* imaging in order to extend the wide-field photometry down to small galactocentric radii, and presented a catalogue of 17620 globular cluster candidates across a radius range

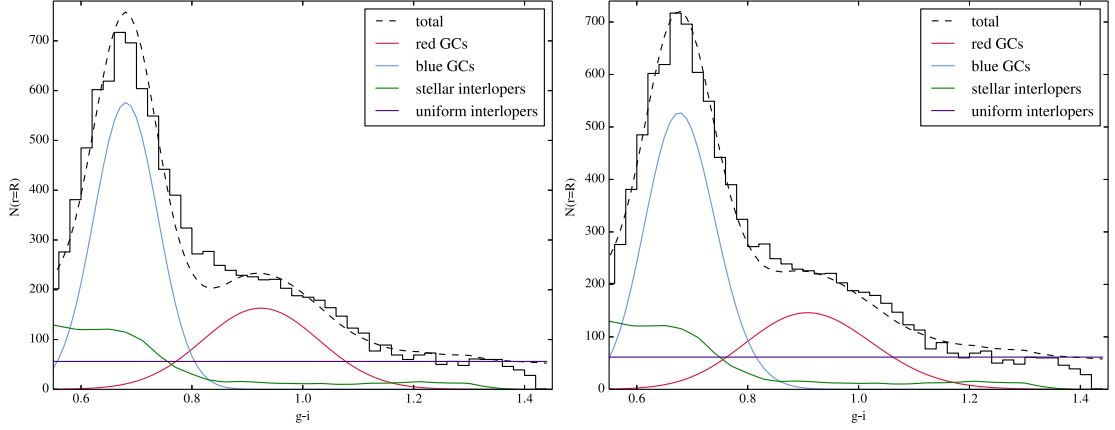


Figure 2.9: Removing the colour bump. Left: the colour distribution of the catalogue (shown by the histogram) is discrepant with the inference from a fixed-peak Gaussian model (shown by the dashed line) at intermediate colours. This motivates a model in which the colour distributions are also a function of radius. Right: allowing for radial gradients in the colour distributions of the globular clusters alleviates the discrepancy between model and data at intermediate colours.

from 1.3 kpc to 445 kpc and to a depth of 24 magnitudes in the g -band. By treating the catalogue as being composed of two globular cluster populations and two contaminant populations, we used a Bayesian framework to infer the colour, luminosity and radial profiles of each.

Our model for the contaminant contribution to the catalogue improves on previous studies by using colours and luminosities that are distributed according to synthetic observations of Milky Way and Sagittarius stars. This is important for distinguishing globular clusters from stars and correctly characterising their radial profiles, as the red-giant branch of the Sagittarius stream lies extremely close to the blue globular cluster population in colour-magnitude space. The use of this model, in conjunction with a uniformly-distributed component to account for other sources of contamination, significantly changes our inference on the globular cluster distributions, and highlights the importance of modelling the interloper populations in a physically-motivated way. We also confirm previous findings of a colour gradient with galactocentric radius within each globular cluster population, and incorporate this in our model using a log-linear relation.

The extensiveness and completeness of this catalogue has allowed us to characterise M87's globular cluster populations in a robust, unbiased way. This makes it an ideal starting point for further studies based on subsamples of these populations, in which selection criteria would otherwise cause the apparent colour and spatial distributions to

deviate from the true ones.

GALAXY STRUCTURE FROM MULTIPLE TRACERS: II. M87 FROM PARSEC TO MEGAPARSEC SCALES

Abstract

Following a number of conflicting studies of M87's mass profile, we undertake a dynamical analysis of multiple tracer populations to constrain its mass over a large radius range. We combine stellar kinematics in the central regions with the dynamics of 612 globular clusters out to 200 kpc and satellite galaxies extending to scales comparable with the virial radius. Using a spherical Jeans analysis, we are able to disentangle the mass contributions from the dark and baryonic components and set constraints on the structure of each. Assuming isotropy, we explore four different models for the dark matter halo and find that a centrally-cored dark matter distribution is preferred. We infer a stellar mass-to-light ratio $Y_{\star,v} = 6.9 \pm 0.1$ – consistent with a Salpeter-like IMF – and a core radius $r_c = 67 \pm 20$ kpc. We then introduce anisotropy and find that, while the halo remains clearly cored, the radial stellar anisotropy has a strong impact on both $Y_{\star,v}$ and the core's radius; here we find $Y_{\star,v} = 3.50^{+0.32}_{-0.36}$ – consistent with a Chabrier-like IMF – and $r_c = 19.00^{+8.38}_{-8.34}$ kpc. Thus the presence of a core at the centre of the dark halo is robust against anisotropy assumptions, while the stellar mass and core size are not. We are able to reconcile previously discrepant studies by showing that modelling the globular cluster data alone leads to the very different inference of a super-NFW cusp, thus highlighting the value of multiple-population modelling, and we point to the possible role of M87's AGN and the cluster environment in forming the central dark matter core.

3.1 Introduction

The Λ CDM paradigm of structure formation has been very successful in describing the Universe on large scales, but there remains some tension regarding its predictions about galaxy structure. In particular, one main prediction of cold, collisionless gravitational collapse is the formation of a central cusp in the density profile of the dark matter haloes that envelope galaxies, with $\rho_{DM} \sim r^{-\gamma}$ and $\gamma = 1$ at small radii (Navarro et al., 1997; Navarro et al., 2010). However, real haloes also contain baryons, and the imprint of baryonic physics on the dark matter distribution could be significant. For instance, feedback from supernovae and AGN, as well as dynamical friction from infalling satellites, could lead to some degree of heating and expansion (e.g. Mashchenko et al., 2006; Laporte et al., 2012; Governato et al., 2012; Velliscig et al., 2014), thus hollowing out the dark matter, whilst the cooling and condensation of baryons could increase the density in the central regions via adiabatic contraction (e.g. Blumenthal et al., 1986; Gnedin et al., 2004). The current observational picture reflects this complexity, with the haloes of an increasing number of dwarf galaxies being found to favour cored or only weakly cuspy central profiles. For instance, the recent local surveys THINGS and LITTLE THINGS (Hunter et al., 2007) found a large fraction of dwarf field galaxies to have dark matter density profiles that go as $\rho \sim r^{-0.4}$ within the central kiloparsec, and studies of low-surface-brightness galaxies also point to relatively flat central profiles with a large scatter (e.g. de Naray & Spekkens, 2011).

Whilst a great deal of progress has been made in constraining the halo structure of low-surface-brightness galaxies and dwarf spheroids, where dark matter dominates over the baryonic mass, the situation is much more complicated for their massive elliptical counterparts. Here, our ignorance about the IMF introduces a degeneracy between dark and luminous matter, which makes it hard to constrain the behaviour of the dark matter in the inner regions; equally, the task of probing the gravitational potential at large radii is made hard by the fact that their outskirts are notoriously faint. However, one way of significantly alleviating these degeneracies is to use multiple dynamical tracer populations, spanning a range of galactocentric radii (e.g. Schuberth et al., 2010; Walker & Peñarrubia, 2011; Napolitano et al., 2014). Indeed, massive elliptical galaxies are often home to large populations of planetary nebulae, globular clusters and even, in the case of BCGs, satellite galaxies, and each of these populations, with its own signature spatial distribution and kinematic profile, can be used as an independent probe of the gravitational potential. The pool of ETGs for which such an analysis has been carried out is currently too small for any meaningful conclusions to be drawn, though early evidence also suggests that the NFW

profile may not provide a good universal fit: studies of BCGs by (Newman et al., 2013, superseding the earlier analyses by Sand et al., 2002; Sand et al., 2004; Sand et al., 2008; Newman et al., 2011) have found evidence for sub-NFW density profiles in clusters, while a handful of studies of field ellipticals (e.g. Grillo, 2012; Sonnenfeld et al., 2012), have found super-NFW densities. Clearly, a lot remains to be understood here, from both observational and theoretical perspectives.

The massive elliptical M87, located at the centre of the Virgo cluster, is an ideal subject for continuing such studies, as it has an enormous globular cluster population (estimated as $\sim 12,000$, e.g. McLaughlin et al., 1994; Tamura et al., 2006b and Chapter 2), making it one of the richest globular cluster hosts in the local Universe. Further, the sample of this population for which we have high-resolution kinematic data has been greatly expanded in recent years through the wide-field study of Strader et al. (2011). Furthermore, the fact that it is a slow rotator and nearly spherical (e.g. Cappellari et al., 2006) makes it suitable for mass modelling under the assumption of spherical symmetry (though see also Emsellem et al., 2014, for recent evidence for kinematic asymmetries). However, the two most recent studies of M87’s mass distribution, both primarily based on Strader’s globular cluster dataset, are markedly inconsistent. The first of these, Agnello et al. (2014), divided the globular cluster sample into three independent populations and used a virial analysis to infer a very cuspy central density profile ($\gamma = 1.57$), while Zhu et al. (2014) combined SAURON central stellar kinematics with the Strader globular cluster data (along with an additional globular cluster sample from Hectospec) and modelled the density profile using a logarithmic potential, thus imposing a core, which they infer to be $r_s = 42 \pm 10$ kpc. Their total inferred stellar masses also differ by almost a factor of two. As M87 is one of the nearest and most well-observed BCGs, it seems unsatisfactory that its mass distribution should still be so poorly constrained, and clearly there remains work to be done. The aim of this chapter is therefore to use a synthesis of globular cluster, satellite and stellar kinematic data in conjunction with flexible mass models in order to infer a density profile which is free to be cuspy or cored as the dynamics dictate.

The chapter is organised as follows: in Section 4.2, we introduce the tracer populations used in our analysis and the associated datasets. In Section 3.3 we describe our mass model and Jeans analysis, the results of which are presented in Section 4.3. We discuss the implications of our findings in Section 4.5, and use Section 4.6 to summarise our main conclusions.

	radial coverage	data type	instruments	sources
stars	1.6×10^{-3} - 210 kpc	photometry	multiple	Kormendy et al. (2009)
	0 - 2.5 kpc	kinematics	SAURON/William Herschel Telescope	Emsellem et al. (2004)
	0 - 0.17 kpc	kinematics	NIFS/Gemini Telescope	Gebhardt et al. (2011)
globular clusters	1.3 - 240 kpc	photometry	MegaPrime/CFHT	Oldham & Auger (2016a)
	2 - 200 kpc	kinematics	multiple	Strader et al. (2011)
satellite galaxies	35 - 1000 kpc	kinematics	multiple (mostly SDSS)	Blakeslee et al. (2009), Kim et al. (2014)

Table 3.1: Various sources of photometric and spectroscopic data for the different tracer populations used in the dynamical analysis.

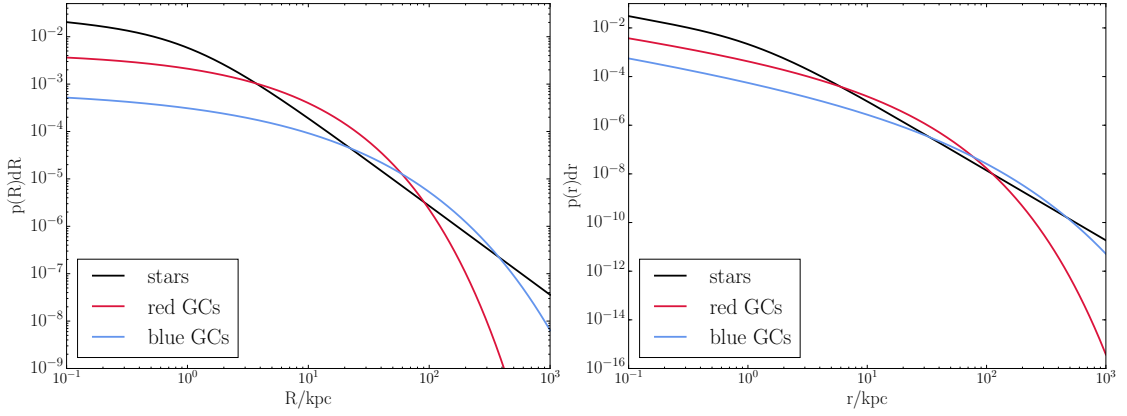


Figure 3.1: Left: Normalised surface brightness profiles for the starlight and the red and blue globular clusters, scaled arbitrarily. The stellar surface brightness was obtained from a fit to the V-band profile presented in Kormendy et al. (2009), while the globular cluster distributions come from the modelling in Chapter 2. Right: Normalised 3D luminosity density profiles for the same three populations, deprojected assuming spherical symmetry.

3.2 Data

To constrain M87's density profile across a wide radius range, we use multiple dynamical tracers, combining stellar kinematics in the central regions with globular cluster dynamics at large radii and satellite galaxies on cluster scales. We can then solve the Jeans equation for each population separately, provided that the underlying density distribution is known. We therefore take data from a number of sources, as summarised in Table 1.

3.2.1 Stars

The deprojected stellar surface brightness profile comes into our analysis at two points: first, the use of the stars as dynamical tracers in the Jeans equation requires us to know their 3D density distribution; second, our goal is to model M87's mass as the sum of dark and luminous components, and the latter is simply the product of the integrated 3D luminosity density with some constant mass-to-light ratio, Υ_* , which can be inferred from the data. We use the radial profile for M87 presented in Kormendy et al. (2009), in which 20 sets of observations across a range of radii and filter systems were synthesised into a single profile in the V band. As M87 is known to have a very extended cD envelope in addition to a stellar core (eg. Chakrabarty, 2007), its surface brightness profile cannot be accurately modelled by a Sérsic profile at both small and large radii. We therefore chose to model it using a more flexible Nuker profile, according to the following relation

$$(3.1) \quad I_*(R) = I_{*,0} \left(\frac{r}{r_b} \right)^{-\zeta_*} \left(1 + \left[\frac{r}{r_b} \right]^{\alpha_*} \right)^{\frac{\zeta_* - \eta_*}{\alpha_*}}$$

with amplitude $I_{*,0}$, break radius $r_b = 1.05$ kpc, inner slope $\zeta_* = 0.186$, outer slope $\eta_* = 1.88$ and break softening $\alpha_* = 1.27$, as shown in Figure 3.2. Though this has the disadvantage of having no analytic deprojection or normalisation, it is much more flexible than a Sérsic profile as it allows both the inner and outer slopes greater freedom. Assuming spherical symmetry, we deproject this profile to give the 3D density shown in Figure 3.1.

The kinematics of the inner $33'' \times 41''$ of M87 have been observed with the integral-field unit (IFU) SAURON (Bacon et al. , 2001), and a catalogue of the first four moments of the Gauss-Hermite expansion of the line-of-sight velocity distribution is available online¹. We use the velocity dispersions, which were obtained from the spectra using a direct pixel fitting routine (Emsellem et al., 2004; Cappellari et. al., 2011). The spectra were adaptively binned to ensure a signal-to-noise of at least 60 per spectral resolution element; uncertainties are generally less than $\sim 20 \text{ kms}^{-1}$, with a mean uncertainty of $\sim 9 \text{ kms}^{-1}$.

M87 is known to host a black hole of mass $\sim 6.6 \times 10^9 M_\odot$ (Gebhardt & Thomas, 2009; Gebhardt et al., 2011), and this should make a significant contribution to stellar velocity dispersions at the smallest radii. The SAURON dataset extends right down to the centre, though its resolution of $1''$ is too low to be able to set constraints on the black hole mass. As we are mainly interested in the behaviour of the dark matter and stellar components, one option to deal with this would be to simply exclude the apertures within the central $\sim 3''$ from our analysis. However, though the contribution of the black hole to the enclosed mass

¹<http://www.strw.leidenuniv.nl/sauron/>

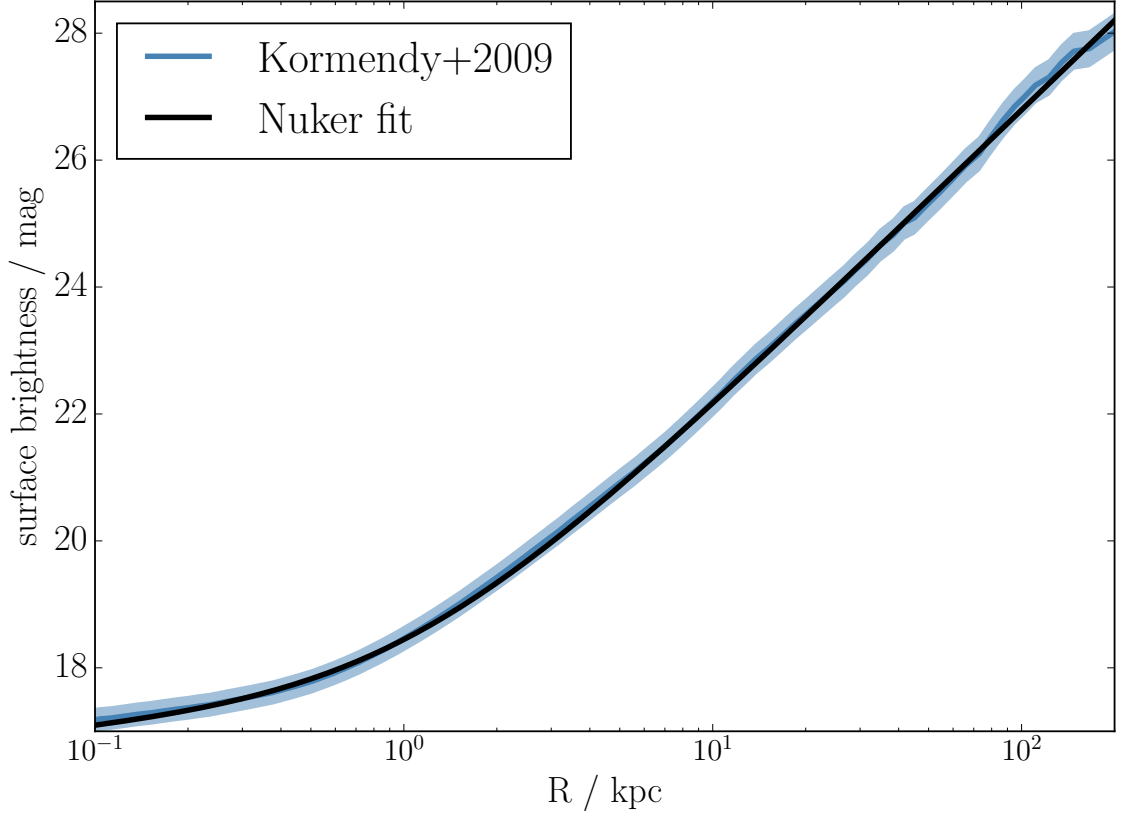


Figure 3.2: The surface brightness profile of M87’s stars has been compiled by Kormendy et al. (2009) (blue curve), and is well described by a Nuker profile (black line) at both small and large radii.

becomes sub-dominant beyond this approximate radius, it still contributes non-negligibly at larger radii and may be covariant with the stellar mass. We therefore include the black hole in our mass model and constrain it using high-resolution kinematics of the central $2''$, as observed with the IFU NIFS on the Gemini North Telescope (Gebhardt et al., 2011). As explained in that paper, these data were obtained from spectra which used laser adaptive optics corrections, and have a resolution of $0.08''$ and a signal-to-noise generally greater than 50. The velocity moments are provided in radius and position angle bins, though our simplifying assumption of spherical symmetry allows us to combine bins azimuthally.

3.2.2 Globular clusters

We use the colour and radial profiles for the globular cluster populations of Chapter 2, as shown in Figure 3.1. The details of the inference are explained fully in that chapter, but in

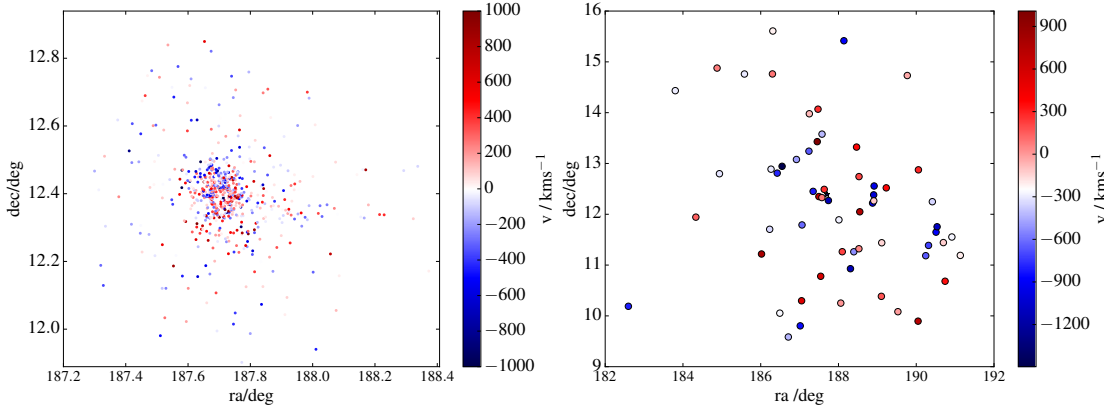


Figure 3.3: Maps of the tracer populations, coloured by velocity with respect to M87. Left: kinematic globular cluster sample. Note that this does not follow the distribution of the underlying population, which is characterised independently using the photometric catalogue of Chapter 2. Right: satellite galaxy sample. M87 itself is plotted as a navy star. Note that we do not expect this spectroscopically-selected sample to be complete.

brief, archival CFHT/MegaPrime images in the *ugriz* bands were used to compile a sample of 17620 globular cluster candidates, selected according to their colours, magnitudes and apparent sizes, and the resulting catalogue was modelled to infer the radial and colour distributions and the relative fractions of the two globular cluster populations. The surface density for each globular cluster component was modelled as a Sérsic profile

$$(3.2) \quad N(R) = N_0 \exp \left[-k_n \left(\frac{R}{R_e} \right)^{\frac{1}{n}} \right]$$

with $k_n = 2n - 0.324$, and with the profiles normalised such that $\int_0^{R_{max}} N(R) dR = 1$, where $R_{max} = 240$ kpc is the radius of the outermost globular cluster in the catalogue. The $g-r$, $g-i$ colours and the globular cluster luminosity function for each globular cluster population were modelled as Gaussians, with radius-dependent colour gradients in all but the latter. The 3D deprojected profiles are also shown in Figure 3.1.

Strader et al. (2011) present a spectroscopic catalogue of 737 globular cluster candidates around M87 at radii from 2 kpc to 200 kpc (plus one object at 800 kpc, which we exclude because it lies outside the region over which our model from the photometry is strictly valid). The catalogue combines new measurements for 451 globular clusters – obtained using Keck/DEIMOS (Faber et al., 2003), Keck/LRIS (Oke et al., 1995) and MMT/Hectospec (Fabricant et al., 2005) – with literature data, and provides a classification of objects as blue globular clusters, red globular clusters, ultra-compact dwarfs and transient/unknown along

with SDSS-band $g - r$ and $g - i$ colours. We cross-correlate our photometric catalogue with this spectroscopic catalogue, selecting only the objects classified in Strader et al. (2011) as globular clusters, to obtain a sample of 612 globular clusters with complete luminosity, spatial, colour and kinematic information. In the analysis that follows, we choose to use the photometry of Chapter 2 over that provided in Strader et al. (2011), for consistency with our globular cluster colour distributions.

3.2.3 Satellite galaxies

The Extended Virgo Cluster Catalogue (EVCC, Kim et al., 2014) provides the redshifts and positions on the sky of 1589 galaxies in a footprint of 725 deg^2 centred on M87, extending to 3.5 times the cluster virial radius. The redshifts are compiled from the SDSS DR7 release and the NASA Extragalactic Database (NED), and each object is classified as either a certain cluster member, a possible member or a background source based on morphological and spectroscopic criteria. As it is important that our sample only contains galaxies moving in M87's halo potential, we selected only those objects classed as certain members according to both criteria, and we further cross-correlated these with the catalogue of Blakeslee et al. (2009), which used surface brightness fluctuations to calculate distance moduli. To avoid contamination from the W cloud, a slightly more distant component of the Virgo cluster at a characteristic distance of 23 Mpc, we imposed a distance cut of 20 Mpc; further, to separate the A cloud (centred on M87) from the B cloud (centred on M49) we also imposed a declination angle cut of 9.5 degrees and a radius cut of 1 Mpc. The spatial and velocity distributions of the resulting sample are shown in Figure 3.3: it comprises 60 galaxies, with radii relative to M87 ranging from 35 kpc to 1 Mpc.

We could use this tracer population in the same way as the stars and the globular clusters and require its velocity dispersion profile to satisfy the Jeans equation so as to obtain a further probe of the mass on scales comparable to the virial radius. However, as noted earlier, the Jeans equation requires us to know the tracer density distribution and, as our satellite sample is most likely highly incomplete, we do not have access to this quantity. Most importantly, the sample of satellites we use has been selected spectroscopically, and this imposes a non-trivial selection function which may alter the spatial distribution from the true underlying one, leading us to draw incorrect conclusions from a Jeans analysis (see Chapter 2, Introduction). Instead, we choose to use these satellites to give an estimate of the total mass at large radii and so give a further constraint in our inference on the mass profile. We do this using the virial mass estimator developed in Watkins et al. (2010), which is designed to be robust against simple approximations to the true distributions.

Specifically, we use the estimator of the enclosed 3D mass within a radius r_{out} given projected positions R and line-of-sight velocities v_{los} :

We use the mass estimator given by

$$(3.3) \quad M(r < r_{out}) = \frac{C}{G} \langle v_{los}^2 R^\mu \rangle,$$

where

$$(3.4) \quad C = \frac{\mu + \nu - 2\beta}{I_{\mu,\nu}} r_{out}^{1-\mu}$$

and

$$(3.5) \quad I_{\mu,\nu} = \frac{\pi^{0.5} \Gamma(\frac{\mu}{2} + 1)}{4 \Gamma(\frac{\mu}{2} + \frac{5}{2})} [\mu + 3 - \beta(\mu + 2)].$$

Here β is the anisotropy parameter

$$(3.6) \quad \beta = 1 - \frac{\sigma_t^2}{\sigma_r^2}$$

for radial and tangential velocity dispersions σ_r^2 and σ_t^2 ; μ and ν are the slopes of the potential and tracer density respectively, both assumed to be scale-free. As mentioned previously, our satellite sample is likely to be incomplete, and this means we cannot infer β , μ and ν directly from the data; instead we calibrate their values using simulations. Following Deason et al. (2013), we use the $z = 0$ halo catalogue of the first MultiDark simulation, described in detail in Prada et al. (2012). This uses the WMAP5 cosmology and contains about 8.6 billion particles per Gpc/h^3 : the halo finder uses the bound density maximum technique described in Klypin & Holtzman (1997). We identify all haloes with more than 30 subhaloes and treat each subhalo as a distinct satellite galaxy. We then use the subhalo velocities and positions and the parent halo mass profiles to infer the posterior distributions on β , μ and ν that best describe the global properties of the population, and use these to generate a posterior on the total mass $M(r < r_{out})$ of our cluster, whose median and standard deviation we use to re-select haloes from the simulation and iterate the procedure until our the inference on μ , β and ν converges. We report median values $\beta = 0.30$, $\mu = 0.15$, $\nu = 1.9$, and find a cluster mass $\log(M(r < 985 \text{kpc})/M_\odot) = 14.11 \pm 0.19$.

To test the calibration, we use the median values of β , μ and ν to apply the virial mass estimator to the simulated parent haloes, and show the resulting comparison in Figure 3.4.

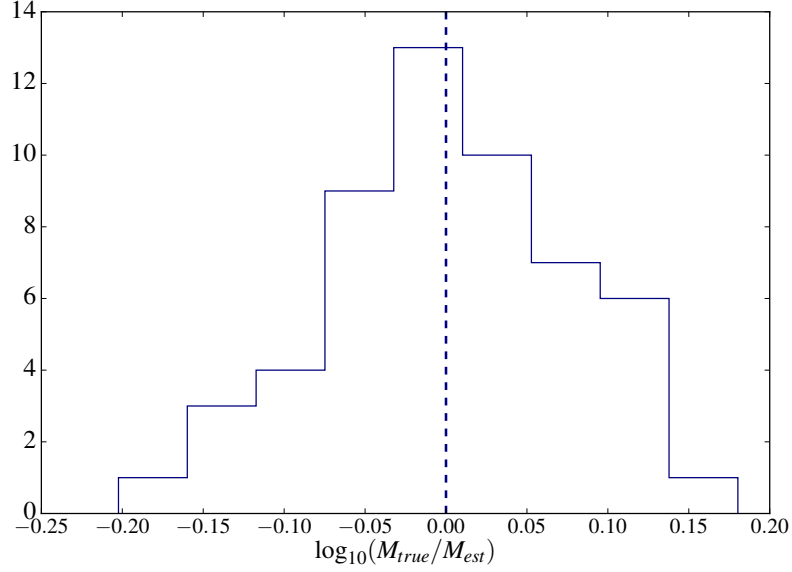


Figure 3.4: The ratio of the true halo mass to that estimated from the virial mass estimator, as calibrated from the simulations. The spread in estimates reflects the fact that the parameters β , μ and ν do in fact vary between the different groups, and are not truly global as we have assumed. However, the scatter is small at 0.1 dex and the median offset is negligible, indicating the robustness of the estimator.

Encouragingly, the masses are consistent, with a negligible median offset and a scatter of 0.1 dex, thus illustrating the robustness of the estimator. The applicability of this to M87 is then dependent on the assumption that the properties of the simulated galaxy satellite populations are representative of real galaxies.

3.3 Modelling

3.3.1 Mass models

We model the galaxy density profile as the sum of luminous and dark matter components plus a black hole:

$$(3.7) \quad \rho(r) = \rho_{\star}(r) + \rho_{DM}(r) + \rho_{BH}(r).$$

The stellar density profile is obtained by deprojecting the stellar surface brightness and scaling by the stellar mass-to-light ratio Υ_{\star} , which we assume to be constant with radius

and include as a free parameter in the model. To allow flexibility in the dark matter density profile and to explore the impact of changing the halo model on the mass inference, we carry out our analysis using four different halo profiles. The first is a standard NFW profile,

$$(3.8) \quad \rho_{DM}(r) = \frac{\rho_0}{4\pi} \left(\frac{r}{r_s} \right)^{-1} \left(1 + \frac{r}{r_s} \right)^{-2},$$

with scale radius r_s and density scale ρ_0 , and cumulative mass

$$(3.9) \quad M_{DM}(r) = \rho_0 r_s^3 \left(\ln \left(1 + \frac{r}{r_s} \right) - \frac{r}{r + r_s} \right).$$

Following Zhu et al. (2014), we also use a cored isothermal profile with a logarithmic potential (the LOG model)

$$(3.10) \quad \rho_{DM}(r) = \frac{\rho_0 r_s^2}{4\pi} \frac{3r_s^2 + r^2}{(r_s^2 + r^2)^2}.$$

This model is inherently cored, but in contrast to the NFW has a $\rho \sim r^{-2}$ dependence at large radii: for a given core, then, this profile allows more dark matter to be placed at large distances from the galaxy centre. Its cumulative mass is given by

$$(3.11) \quad M_{DM}(r) = \frac{\rho_0 r_s^2 r^3}{r_s^2 + r^2}.$$

We also use a generalised NFW (gNFW),

$$(3.12) \quad \rho_{DM}(r) = \frac{\rho_0}{4\pi} \left(\frac{r}{r_s} \right)^{-\gamma} \left(1 + \frac{r}{r_s} \right)^{\gamma-3},$$

with free parameters γ, ρ_0 and r_s , where γ is the inner slope. This has the advantage of leaving the profile free to choose between cusps ($\gamma \geq 1$) and cores ($\gamma = 0$) in the centre, while still becoming NFW-like at large radii in a way that is consistent with both simulations and observations. This gives a cumulative mass profile that can be related to the Gauss hypergeometric function ${}_2F_1$ as

$$(3.13) \quad M_{DM}(r) = \frac{\rho_0 r_s^3}{\omega} \left(\frac{r}{r_s} \right)^\omega {}_2F_1[\omega, \omega; 1 + \omega; -\frac{r}{r_s}]$$

where $\omega = 3 - \gamma$. The final profile we use is a cored generalised NFW (cgNFW)

$$(3.14) \quad \rho_{DM}(r) = \rho_0 \left(\frac{r+r_c}{r_s} \right)^{-\gamma} \left(1 + \frac{r}{r_s} \right)^{\gamma-3}.$$

in which r_c is now the scale radius of the core. This is a more general case of the gNFW, and, in addition to allowing the data to choose between cusps ($r_c = 0$, $\gamma = 1$) and cores ($r_c > 0$, or $r_c = 0$ and $\gamma = 0$), it has additional flexibility at intermediate radii. We carry out the integration of the cumulative mass for this profile numerically. We note, moreover, that the NFW, gNFW and cgNFW profiles form a family of nested models. Furthermore, since X-ray observations show that M87 is situated at the centre of the potential well of Virgo (e.g. O’Sullivan et al., 2001), we assume that there is no mass contribution from other dark haloes in the cluster.

The black hole is simply a point mass at the origin,

$$(3.15) \quad \rho_{BH}(r) = \frac{M_{BH}}{4\pi r^2} \delta(r),$$

giving a constant term in the cumulative mass distribution.

As the SAURON dataset far outweighs the globular cluster, NIFS and satellite datasets in terms of size, we regularise its contribution to the likelihood calculation by additionally fitting for a ‘noise’ parameter Δ_{σ_*} . This can be interpreted as accounting for scatter in the data not included in the uncertainties or, alternatively, as modifying the relative weight given to these data, and is added in quadrature to the measured uncertainties on the velocity dispersion.

Our overall model therefore has a number of free parameters dependent on the halo model and our assumptions about the anisotropy. In the isotropic case, which we treat first, the number of parameters varies between five and seven. In common for all halo models are the mass-to-light ratio Υ_* , the normalisation of the dark matter halo $\log(\rho_0)$, the scale radius r_s , the black hole mass $\log(M_{BH})$ and the noise in the SAURON data, $\Delta_{\sigma_*}^2$. The gNFW and cgNFW halo models then add the free parameters (γ) , (γ, r_c) respectively. When the orbital anisotropy of each tracer population is allowed to vary, this adds three parameters, as will be discussed in Section 3.3.2. In our notation in the following sections, we use the gNFW parameter set whenever we write the model parameters explicitly, but this should be understood as standing in for any of the halo models.

3.3.2 Jeans modelling

Given the stellar velocity dispersion and the globular cluster and satellite velocities, we want to infer the posterior probability distribution on M87’s density profile. For the satellite

galaxies, this involves a direct comparison of the mass calculated from our virial estimator and that obtained by integrating the proposed density profile. For the stars and globular clusters, on the other hand, we can relate the observed velocities and velocity dispersions to the density profile via the Jeans equation. Assuming spherical symmetry and dynamical equilibrium, the Jeans equation has the simple form

$$(3.16) \quad \frac{d}{dr}(l\sigma_r^2) + 2\frac{\beta(r)}{r}l\sigma_r^2 = l(r)\frac{GM(r)}{r^2}$$

where $l(r)$ is the luminosity density of the tracer, $\sigma_r(r)$ the radial velocity dispersion and $\beta(r)$ is the anisotropy parameter defined in Equation 3.6. This is a first-order differential equation with general solution

$$(3.17) \quad l(r)\sigma_r^2(r) = \frac{1}{f(r)} \int_r^\infty f(s)l(s)\frac{GM(s)}{s^2}ds$$

where

$$(3.18) \quad f(r) = f(r_i) \exp \left[\int_{r_i}^r 2\beta(s)\frac{ds}{s} \right]$$

(e.g. van der Marel, 1994; Mamon & Łokas, 2005). Projecting this along the line of sight gives the line-of-sight velocity dispersion, $\sigma_{los}(R)$, as

$$(3.19) \quad \frac{1}{2G}I(R)\sigma_{los}(R)^2 = \int_R^\infty \frac{l\sigma_r^2 r dr}{\sqrt{r^2 - R^2}} - R^2 \int_R^\infty \frac{\beta l\sigma_r^2 dr}{r\sqrt{r^2 - R^2}}$$

which, for certain choices of anisotropy parameterisations, can be written in the simple form

$$(3.20) \quad I(R)\sigma_{los}^2(R) = 2G \int_R^\infty lK_\beta M \frac{dr}{r}$$

with the kernel K_β dependent on the particular anisotropy model. Initially, we assume all orbits to be isotropic; later we consider the effect of anisotropy on our inference by modelling the stellar population as following a radially-dependent anistropy profile and each globular cluster population with a constant, non-zero anisotropy. For the stars, we use a scaled Osipkov-Merritt (Osipkov, 1979; Merritt, 1985) profile

$$(3.21) \quad \beta(r) = \beta_\infty \frac{r^2}{r^2 + r_a^2};$$

which is centrally isotropic and becomes radially anisotropic at large radii, tending to the asymptotic anisotropy β_∞ . The kernel for the scaled Osipkov-Merritt profile is presented in the Appendix; the kernel for the constant-anisotropy case is given in Mamon & Łokas (2005), and we refer the reader there for further details.

Given a prescription for the anisotropy, the surface brightness and luminosity profiles of Figure 3.1 and the mass model, then, we can calculate line-of-sight velocity dispersions σ_{los} from the Jeans equation and so infer the posterior probability distribution of the model parameters, based on the data in hand.

3.3.3 Statistical analysis

We use a Bayesian approach to infer the posterior probability distribution of our density profile parameters, given the data. Bayes theorem states that the posterior distribution is proportional to the product of the likelihood function of the data given the model and the priors on the model, and so the task here is to construct sensible likelihood functions for each dataset. The total likelihood is then, in turn, the product of these, as each constitutes an independent set of measurements.

First, as we have velocity dispersion measurements for the stars, the likelihood of observing a particular velocity dispersion at radius R is assumed to be Gaussian, with a standard deviation equal to the uncertainty. Thus the k^{th} stellar velocity dispersion measurement gives a contribution to the likelihood:

$$(3.22) \quad \ln L_{\star,k}(\sigma_k, R_k | \vec{M}) = -0.5 \left(\frac{\sigma_k - \sigma_m}{\delta_{\sigma_k}} \right)^2 - 0.5 \ln(2\pi \delta_{\sigma_k}^2)$$

for uncertainty δ_{σ_k} and model prediction σ_m^2 , and model parameters $\vec{M} = (Y_\star, M_{BH}, \Delta_{\sigma_\star}, \rho_0, \gamma, r_s)$. As the observations in each aperture are independent, the total log likelihood of observing the ensemble is just the sum:

$$(3.23) \quad \ln L_\star = \sum_k \ln L_{\star,k}.$$

Note that, for measurement uncertainty $\Delta_{\sigma_k}^2$, the regularisation of the SAURON data gives a total uncertainty $\delta_{\sigma_k}^2 = \Delta_{\sigma_k}^2 + \Delta_{\sigma_\star}^2$. For the NIFS data, on the other hand, $\delta_{\sigma_k}^2 = \Delta_{\sigma_k}^2$.

The virial mass estimate from the satellites can be treated in a similar way, though here we only have one measurement:

$$(3.24) \quad \ln L_{sat} = -0.5 \left(\frac{\log M_{sat} - \log M_{mod}}{\delta M_{sat}} \right)^2 - 0.5 \ln(2\pi \delta M_{sat}^2)$$

for virial mass estimate M_{sat} , model mass M_{mod} and the logarithm of the uncertainty from the mass estimator, δM_{sat} . This is a direct mass comparison since we have already converted the satellite kinematics to a robust mass estimate.

For the globular clusters we can assign probabilities, based on the colour and position information provided in the photometric catalogue, of each globular cluster belonging to either the red or the blue population, although we are not able to classify them with certainty. As the two populations are assumed to be dynamically decoupled, we do not expect their velocity dispersion profiles to be the same. In contrast to the other tracer populations, then, we calculate the likelihood of observing a globular cluster with a particular velocity under the assumption that the velocity distribution of each globular cluster population can be described by a Gaussian with a standard deviation given by the velocity dispersion, such that

$$(3.25) \quad \ln L_{GC,k} = -0.5 \left(\frac{v_k^2}{\delta v_k^2 + \sigma_m^2} \right) - 0.5 \ln \left(2\pi(\delta v_k^2 + \sigma_m^2) \right)$$

where v_k and δv_k are the measured line-of-sight velocity and velocity uncertainty of the k^{th} globular cluster (centered on M87's heliocentric velocity $v_{M87} = 1284 \text{ kms}^{-1}$, Cappellari et. al., 2011), and the line-of-sight velocity dispersion at the location of each globular cluster, σ_m , is modelled separately for the red and the blue populations.

To account for the uncertainty in assigning each globular cluster to either the red or the blue population, for each set of model parameters \vec{M} we draw 1000 Monte Carlo samples which stochastically explore the population distribution based on the colour, magnitude and position information for the individual globular clusters. We then marginalise over the samples to give a final contribution to the likelihood. Whilst some globular clusters have either very high or very low probabilities of belonging to one of the globular cluster populations, with small uncertainty, there also exists a significant fraction with comparable probabilities of belonging to either: for these objects, it would not be meaningful to simply assign them to one population or the other. A further advantage of this stochastic sampling is that it allows us to explore different combinations of red and blue globular clusters.

As each tracer population is independent, the final log-likelihood of any set of model parameters is the sum over all contributions:

$$(3.26) \quad \ln L = \sum_k \ln L_{\star,k} + \sum_k \ln L_{GC,k} + \ln L_{M_{sat}}.$$

We explore the parameter space using the ensemble-sampling code EMCEE (Foreman-Mackey et al., 2013).

halo model	M_{\star}/L	$\log(\rho_{DM})$	r_s	γ	r_c	$\log(M_{vir})$	R_{vir}
NFW	6.6 ± 0.1	6.6 ± 0.1	448 ± 75	—	—	$14.39^{+1.11}_{-0.53}$	1620^{+460}_{-360}
LOG	6.9 ± 0.1	7.7 ± 0.1	48 ± 5	—	—	$14.21^{+0.60}_{-0.37}$	1410^{+240}_{-200}
gNFW	6.9 ± 0.1	8.3 ± 0.1	79 ± 10	< 0.14	—	$14.13^{+0.76}_{-0.44}$	1320^{+270}_{-240}
cgNFW*	6.9 ± 0.1	$5.6^{+1.4}_{-1.2}$	273^{+430}_{-180}	$2.54^{+0.33}_{-1.18}$	63^{+11}_{-14}	$14.16^{+0.67}_{-0.40}$	1350^{+240}_{-200}

Table 3.2: Final inference on the parameters in the different $\beta = 0$ models. We report the median values of our inferred posterior distributions, along with the 16th and 84th percentiles as a measure of our uncertainty. For the inner slope γ in the gNFW and cgNFW models, the 95 % confidence value is given. All quantities are measured in units of solar mass, solar luminosity, kilometers per second and kiloparsecs. *Note that the cgNFW posterior is bimodal due to degeneracies inherent in the profile: see the panel of Figure 3.8.

3.4 Results

Initially, we treat all orbits as isotropic and investigate the uncertainty introduced to the mass inference by the choice of halo model. The inference for the four different halo models are presented in Figures 3.5 - 3.8, and are summarised in Table 3.2 along with the associated virial parameters. We find the results of the gNFW, cgNFW and LOG models to be very similar: the gNFW and cgNFW haloes are both centrally cored with break radii ~ 80 kpc, while the LOG model is inherently cored and has a slightly smaller scale radius $r_s = 48 \pm 5$ kpc (note, however, that these are *not* core radii, and that the scale radius is defined differently in the latter case). All three models also agree closely on the high stellar mass-to-light ratio $\Upsilon_{\star,v} \sim 6.9$ and the virial parameters, with $\log(M_{vir}/M_{\odot}) \sim 14.2$ and $R_{vir} \sim 1.3$ Mpc. Indeed, the best cgNFW model approximately recovers the gNFW solution, with an intermediate slope $\gamma \sim 3$ and both the core and break radii comparable to the gNFW scale radius. Like the gNFW, this represents a solution in which the halo is cored centrally and becomes NFW-like at large radii; unlike the gNFW, it allows more flexibility in the intermediate regions, and the fact that this intermediate slope is found to be close to the NFW value $\gamma \sim 3$ suggests that the NFW profile is an adequate description at intermediate radii as well as large radii.

On the other hand, the NFW model predicts a lower mass-to-light ratio $\Upsilon_{\star,v} = 6.60 \pm 0.05$, though it does match the other models in terms of its virial parameters. The difference at small radii arises because of the NFW profile’s hard-wired cusp, which places more dark matter in the central regions at the expense of stellar mass. In Figure 3.9 we plot the mass profiles inferred in each case, along with the associated rotation curves in order to better highlight the differences between the four models. As can be seen, deviations start to arise at larger radii, where the constraints from our data are weakest. These are mainly due to

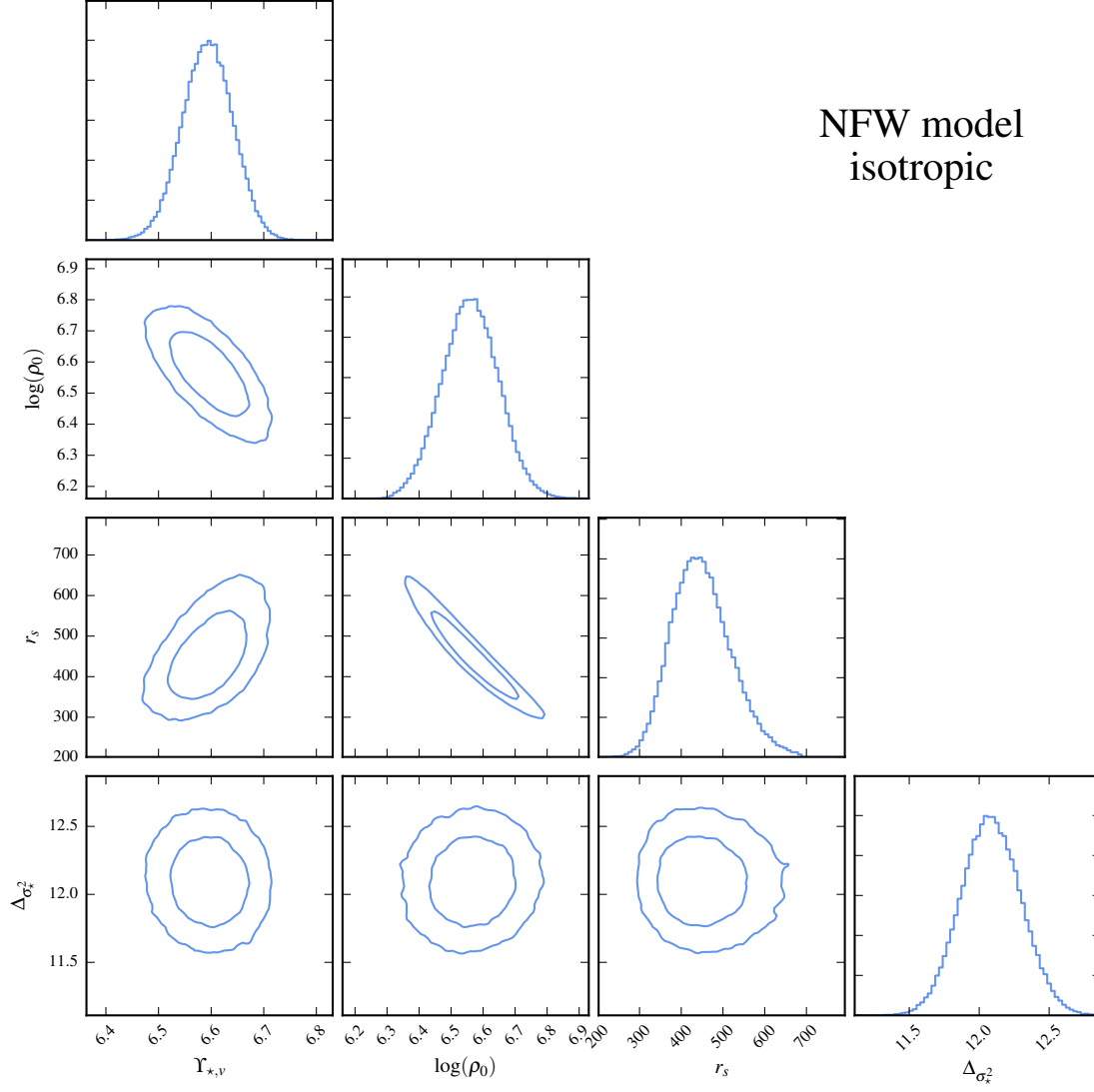


Figure 3.5: Inference on the NFW model parameters, assuming isotropy. This model favours a lower mass-to-light ratio than the others, due to the larger amount of dark matter that the cusp necessarily puts at small radii. As in Tables 2 and 3, all quantities are measured in units of solar mass, solar luminosity, kilometers per second and kiloparsecs.

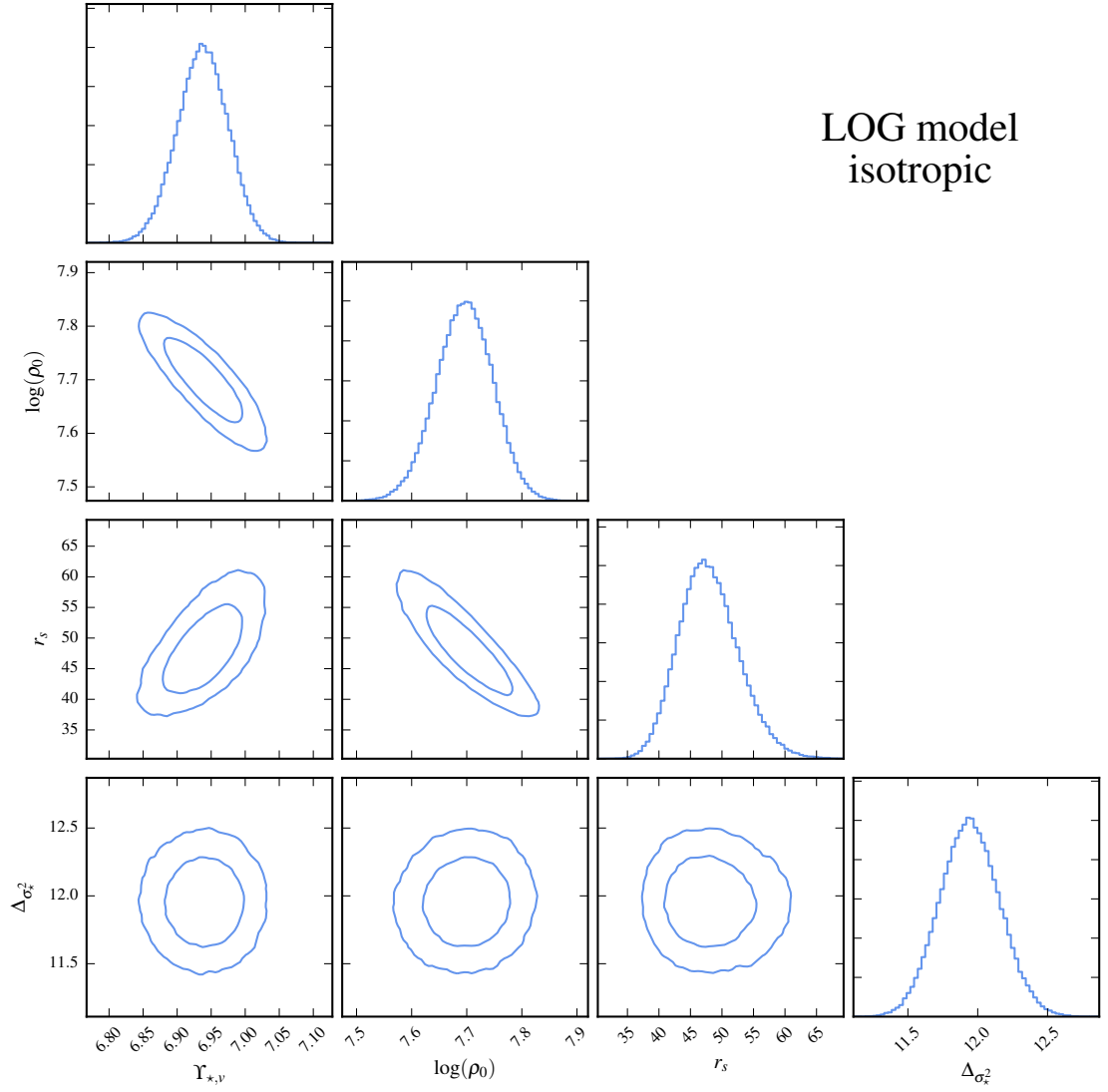


Figure 3.6: Inference on the LOG model parameters, assuming isotropy. All quantities are measured in units of solar mass, solar luminosity, kilometers per second and kiloparsecs.

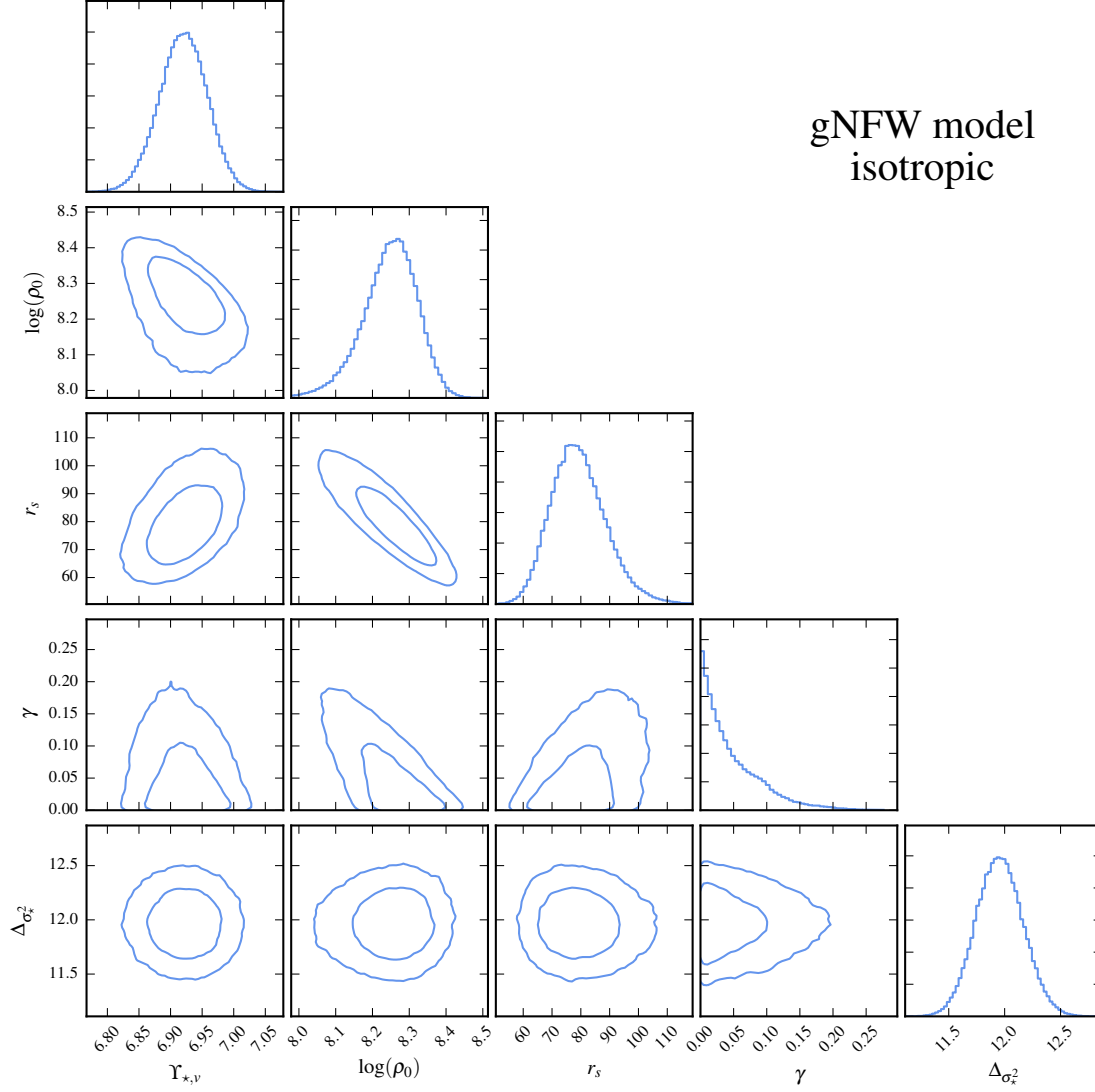


Figure 3.7: Inference on the gNFW model parameters, assuming isotropy. Note that the posterior on the inner slope, γ , hits the lower limit imposed by the prior. All quantities are measured in units of solar mass, solar luminosity, kilometers per second and kiloparsecs.

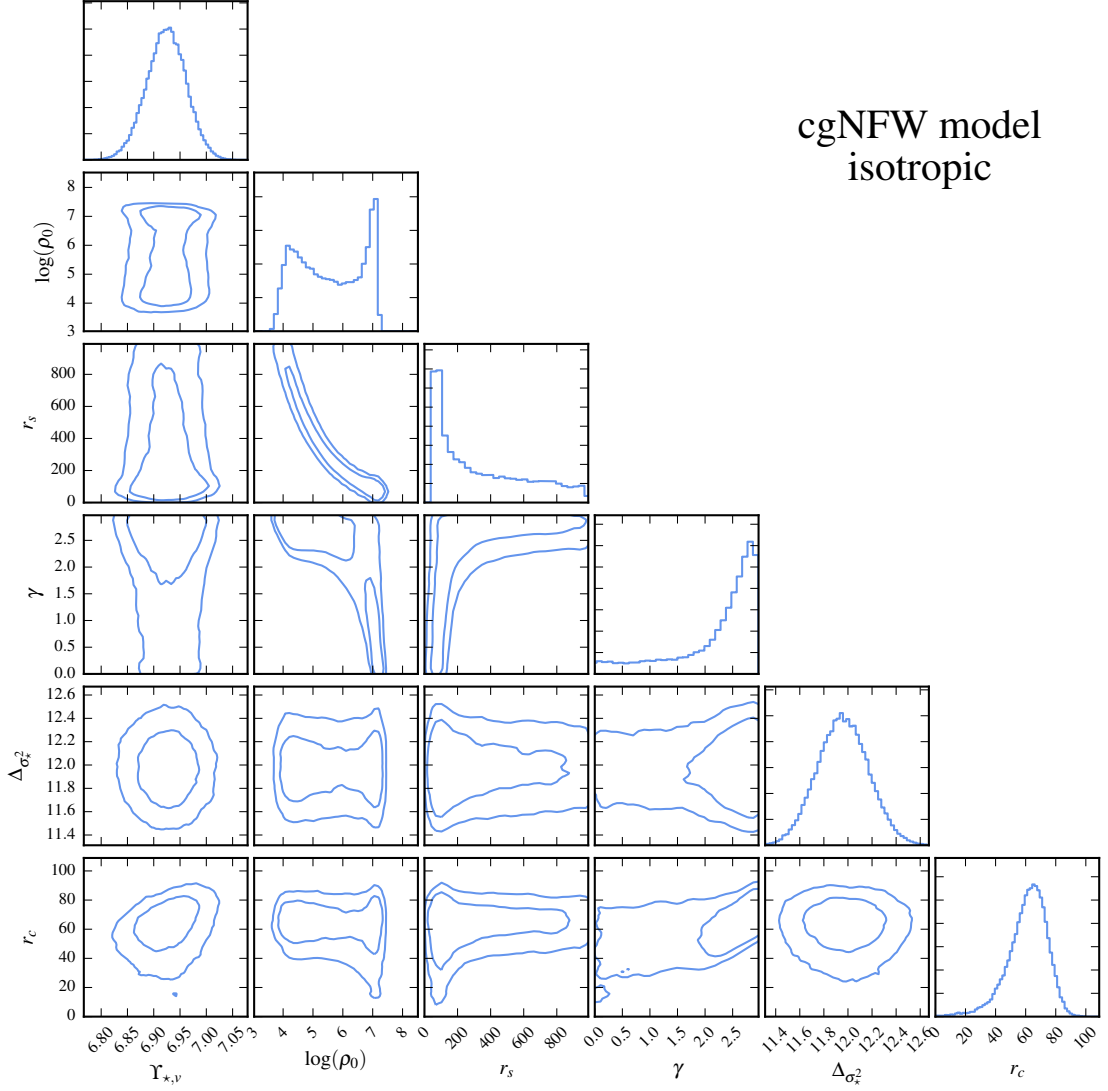


Figure 3.8: Inference on the cgNFW model parameters, assuming isotropy. All quantities are measured in units of solar mass, solar luminosity, kilometers per second and kiloparsecs. Note that our cgNFW modelling implies that the halo is consistent with a cored power law with power law index 3. For the cgNFW profile, this leads to a degeneracy in the parameters γ , r_s and r_c : given a finite core radius, the same solution can be obtained either by having $\gamma \sim 0$ and $r_s \sim r_c$ (in which case ρ_0 is large), or by having $\gamma \sim 3$, in which case r_s (and so ρ_0) is relatively unconstrained. This is what we find here, producing the apparent bimodality in ρ_0 . Both these solutions correspond to the same mass profile.

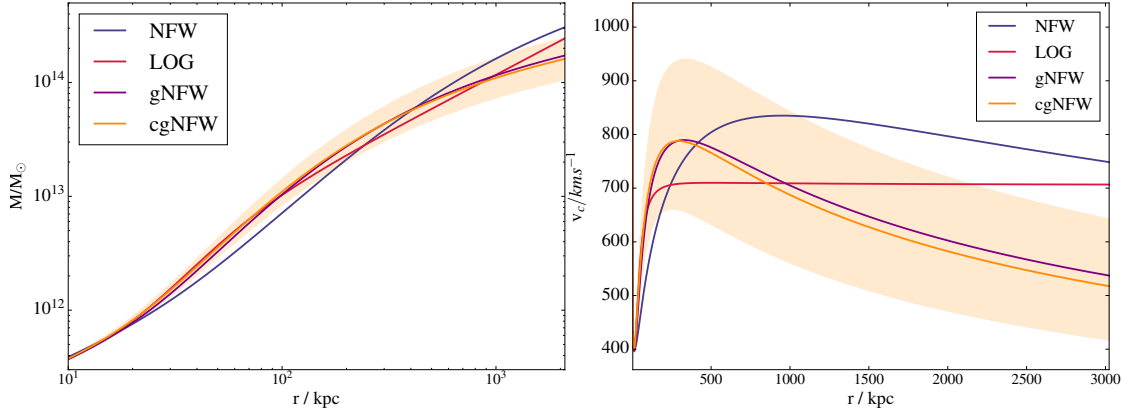


Figure 3.9: Left: mass profiles inferred using the four different halo models, under assumptions of isotropy. At small radii, all are dominated by the stellar mass, so there is little scatter. At larger radii they differ due to the different structures imposed by the halo models: the LOG density profile, for instance, decays as r^{-2} while the others go as r^{-3} , and this places more mass further out in the halo. The NFW is the only profile which does not allow a central core, and its mass profile differs from the other three across a wide radius range. Right: circular velocity curves for the four halo models. This highlights the differences between them. In both figures, the band indicates the 68 % uncertainty region for the cgNFW model.

the different structural features of the models: for instance, the LOG profile decays as r^{-2} at large radii whereas the other three go as r^{-3} , which allows the former to place more mass at large radii. Equally, the NFW profile goes as r^{-1} at small radii whereas the others are either intrinsically cored or have flexible inner slopes (which are found to favour cores here), and this causes the NFW profile to start deviating from the others at smaller radii.

Following this, we test the robustness of our inference on the halo structure against model assumptions by relaxing the isotropy condition and introducing some element of anisotropy into each of the tracer populations. Guided by cosmological N-body simulations, (e.g. Diaferio, 1999; Colín et al., 2000; Wojtak et al., 2005), which find orbits to change smoothly from being radially anisotropic in the galaxy outskirts to being isotropic in the centre – a result of hierarchical formation in which BCGs form by the accretion of infalling satellites which become phase-mixed by the time they reach the galaxy centre – we choose to model the stellar anisotropy using the scaled Osipkov-Merritt profile of Equation 3.21. As the globular cluster data are comparatively sparse, we treat each globular cluster population as having some constant, non-zero anisotropy. This increases our parameter space by four, as we are now also fitting for the scale radius of the stellar scaled Osipkov-Merritt profile r_a , the asymptotic stellar anisotropy β_∞ , and the anisotropies of the two

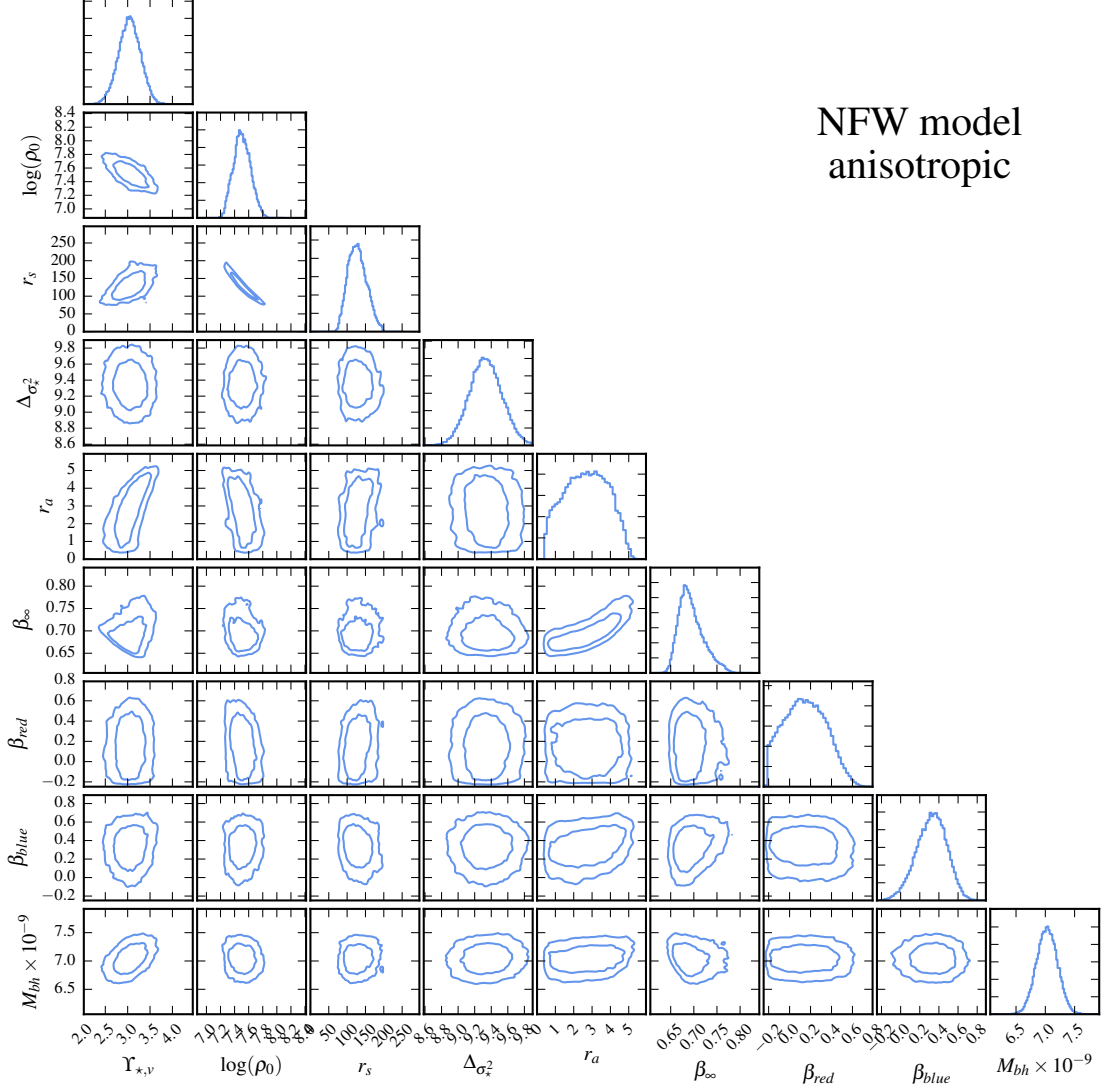


Figure 3.10: Inference on the NFW model parameters, allowing for anisotropy. We characterise the stars with a scaled Osipkov-Merritt profile which becomes isotropic centrally and tends to $\beta = \beta_\infty$ at large radii; for simplicity, each globular cluster population is modelled with constant anisotropy. All quantities are measured in units of solar mass, solar luminosity, kilometers per second and kiloparsecs.

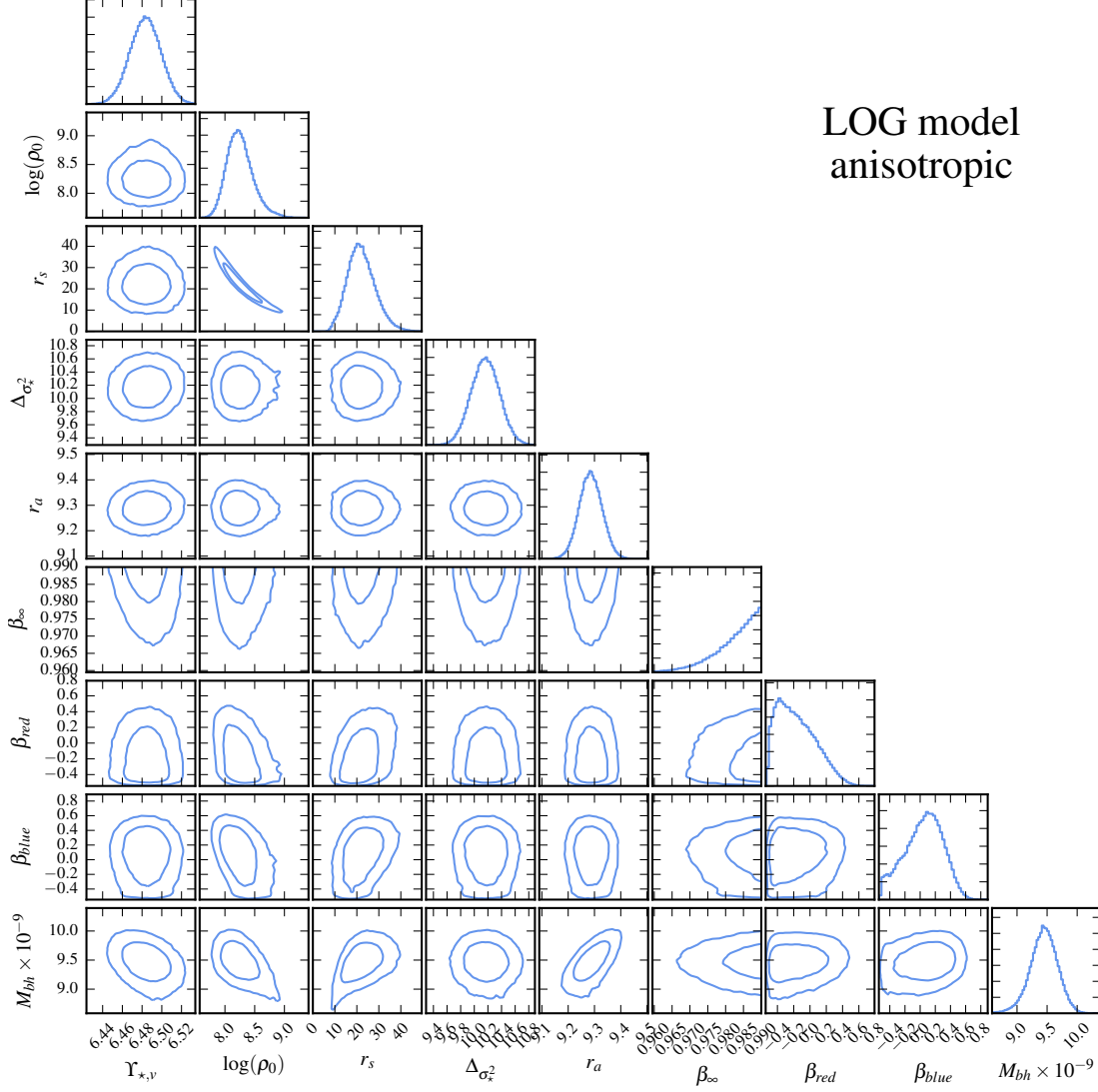


Figure 3.11: Inference on the LOG model parameters, allowing for anisotropy. We characterise the stars with a scaled Osipkov-Merritt profile which becomes isotropic centrally and tends to $\beta = \beta_\infty$ at large radii; for simplicity, each globular cluster population is modelled with constant anisotropy. All quantities are measured in units of solar mass, solar luminosity, kilometers per second and kiloparsecs.

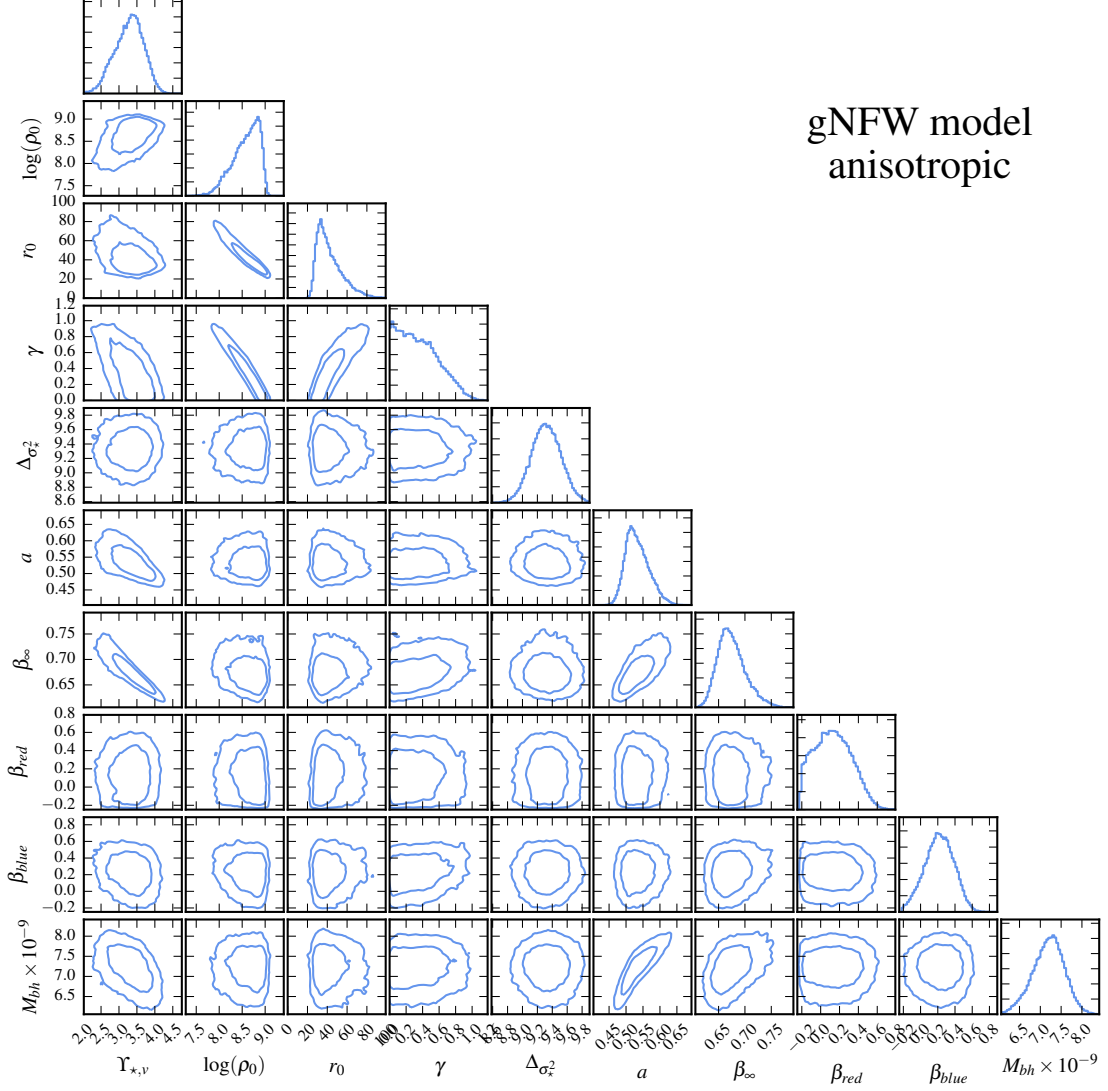


Figure 3.12: Inference on the gNFW model parameters, allowing for anisotropy. We characterise the stars with a scaled Osipkov-Merritt profile which becomes isotropic centrally and tends to $\beta = \beta_\infty$ at large radii; for simplicity, each globular cluster population is modelled with constant anisotropy. All quantities are measured in units of solar mass, solar luminosity, kilometers per second and kiloparsecs.

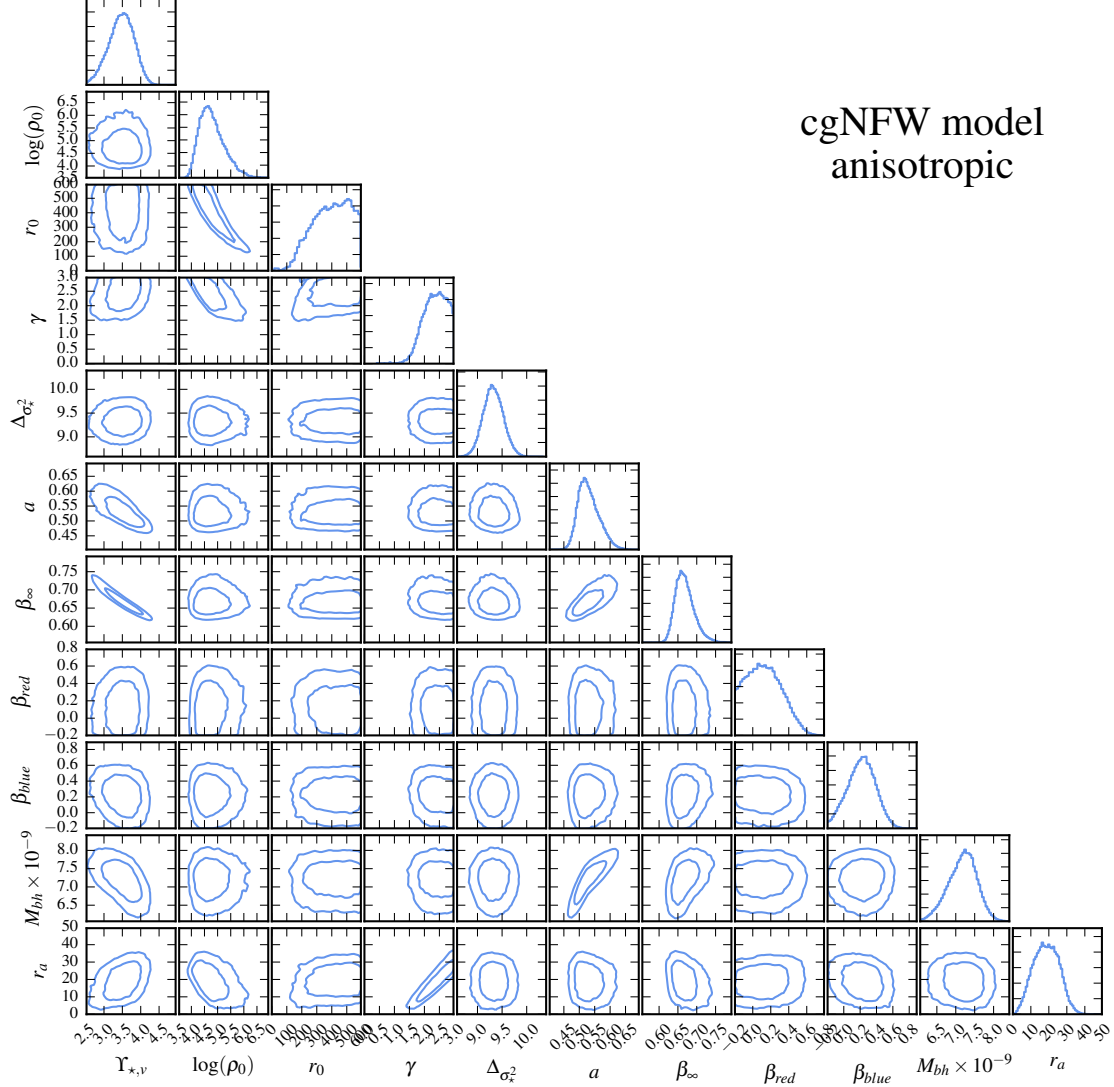


Figure 3.13: Inference on the cgNFW model parameters, allowing for anisotropy. We characterise the stars with a scaled Osipkov-Merritt profile which becomes isotropic centrally and tends to $\beta = \beta_\infty$ at large radii; for simplicity, each globular cluster population is modelled with constant anisotropy. All quantities are measured in units of solar mass, solar luminosity, kilometers per second and kiloparsecs.

globular cluster populations, β_{red} and β_{blue} . Again, we carry out the inference for the four halo models, and present the posteriors in Figures 3.10 - 3.13 and summarise the results in Table 3.3.

Interestingly, we now find some significant deviation among the halo models. Firstly, the LOG profile recovers a model for the stars that is quite isotropic, with $\beta_{\infty} > 0.99$ at the 95% confidence level and $r_a = 9.29 \pm 0.05$ kpc; given that the SAURON data only extend out to ~ 2 kpc, this implies anisotropies $\beta < 0.05$ where we have data (though our data are in projection). This is presumably a result of the hard-wired core. The remaining three models agree, within uncertainties, on the mass profile – though, as might be expected, not so closely as when isotropy is enforced. In these cases, we find that the mass-to-light ratio drops significantly and that the stars become radially anisotropic within scale radii $r_a < 4$ kpc. This covariance between the central mass and anisotropy is a direct manifestation of the mass-anisotropy degeneracy: in the isotropic models, the assumed lack of any radial anisotropy drove the central mass to high values, whereas here we are able to go some way towards breaking the degeneracy through the use of multiple populations tracing the same gravitational potential. Note also that the introduction of radial anisotropy, coupled with the decrease in stellar mass, now allows the dark matter halo to become less cored in the centre: we find a gNFW inner slope which is still distinctly sub-NFW, but less extremely so, with $\gamma < 0.81$ at the 95 % confidence level, and the cgNFW model finds a smaller central core ($r_c = 19.00^{+8.38}_{-8.34}$ kpc) and a sub-NFW intermediate slope $\gamma = 2.39^{+0.39}_{-0.43}$ before transitioning to the r^{-3} regime beyond the scale radius $r_s = 412.1^{+123.0}_{-143.0}$ kpc. For these three profiles, both globular cluster populations are found to have a mild radial anisotropies. We find a slightly higher black hole mass ($M_{BH} = 7.22^{+0.34}_{-0.40}$ for the cgNFW model) than previous estimates (e.g. Gebhardt et al., 2011), though it is consistent within uncertainties. However, we note that this is the first time the black hole mass has been inferred jointly with the halo and stellar mass parameters using the high-spatial-resolution NIFS dataset; as we might anticipate a strong covariance between these parameters – especially between the stellar and black hole masses – a slight change is not too surprising. On the other hand, there are still assumptions in our model which could affect our inference on the black hole mass. For instance, our anisotropy profile enforces isotropy in the centre, which, while a physically sound assumption, restricts the range of parameter space explored by the black hole mass. We are also only making use of the second-order moments of the line-of-sight velocity profile: while this should not cause any systematic bias in the black hole mass inference, we may be able to obtain tighter constraints by incorporating higher-order moments.

We compare the goodness of fit of the four models using a reduced chi-squared criterion,

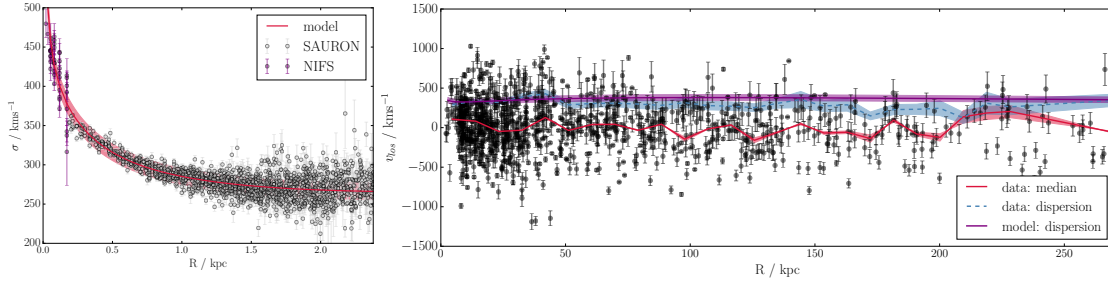


Figure 3.14: Comparison of model and data for M87’s kinematics, for both stars (left) and globular clusters (right). Note that our model does not explicitly fit the binned velocity dispersion of the globular cluster population.

and find that the cgNFW profile provides the best description of the dynamics – noting that the cgNFW, gNFW and NFW are nested models. We therefore present the best cgNFW mass profile in Figure 3.15 (top) and the residuals on the SAURON and NIFS data in Figure 3.15 (bottom). The mass profile appears to be generally consistent with earlier work; the residuals are noticeably better than in the isotropic case, indicating the importance of the anisotropy in the stellar dynamics. We show a comparison of the model and data in Figure 3.14.

We also note that the assumption that the MultiDark simulations are able to accurately reproduce the dynamics of real galaxies may introduce some additional systematic uncertainty into the mass measurement made using the virial estimator in Section 3.2.3. We therefore test the sensitivity of our inference to this uncertainty by carrying out the analysis with twice the calculated uncertainty, and we find that this has a negligible effect on our inferred density model. Essentially, the likelihood is dominated by contributions from the stars and the globular clusters, as the datasets for these populations are far more extensive; the satellite mass estimate rejects models with discrepant extrapolations of the total mass to large radii, but most of the models consistent with the smaller-radius data are still well-behaved at 1 Mpc.

3.5 Discussion

3.5.1 Resolving previous discrepancies

Part of the impetus for this study was to resolve the discrepancies between recent models of M87’s mass structure. The two recent analyses of Zhu et al. (2014) and Agnello et al. (2014), both of which used the Strader et al. (2011) globular cluster kinematics, disagreed on both

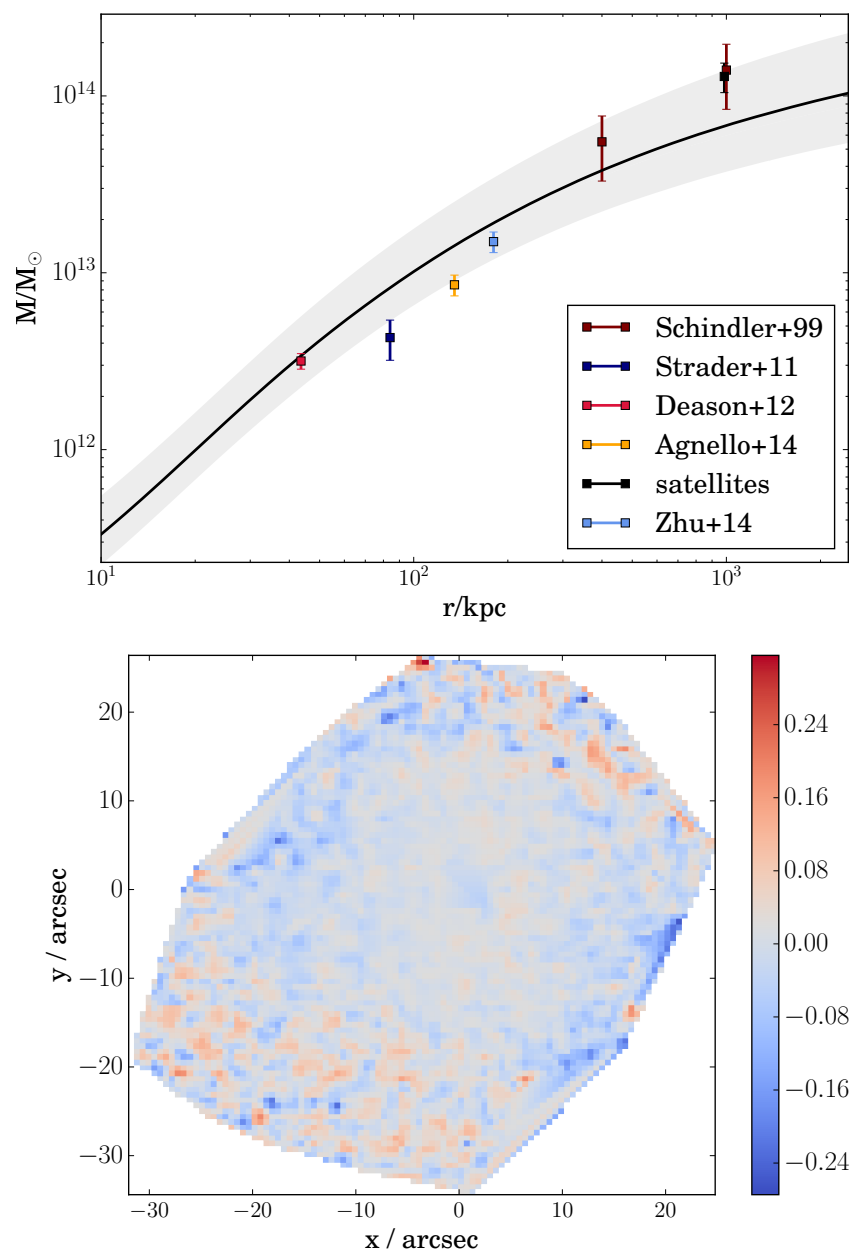


Figure 3.15: Top: Final inference on the mass profile for the anisotropic cgNFW model, with 1σ uncertainties. Overplotted are the mass measurements reported from other studies: these generally bracket our result well. Bottom: Fractional residuals on the stellar velocity dispersion data for the anisotropic cgNFW model.

	M_*/L	$\log(\rho_{DM})$	r_s	γ	$M_{BH} \times 10^{-9}$	r_a	β_∞	β_{red}	β_{blue}	r_c	$\log(M_{vir})$	R_{vir}
NFW	$3.05^{+0.24}_{-0.24}$	$7.51^{+0.13}_{-0.12}$	128^{+28}_{-24}	—	$7.04^{+0.2}_{-0.2}$	$2.78^{+1.26}_{-1.30}$	$0.69^{+0.04}_{0.04}$	$0.16^{+0.21}_{-0.21}$	$0.34^{+0.15}_{-0.17}$	—	$14.00^{+0.37}_{-0.32}$	1190^{+240}_{-120}
LOG	$6.48^{+0.02}_{-0.02}$	$8.25^{+0.23}_{-0.19}$	$22.1^{+6.7}_{-5.7}$	—	$9.5^{+0.22}_{-0.23}$	$9.29^{+0.05}_{-0.05}$	> 0.99	$-0.17^{+0.23}_{-0.21}$	$0.08^{+0.23}_{-0.29}$	—	$14.03^{+0.44}_{-0.29}$	1200^{+250}_{-190}
gNFW	$3.33^{+0.38}_{-0.45}$	$8.64^{+0.24}_{-0.35}$	$40.3^{+16.4}_{-8.8}$	< 0.81	$7.2^{+0.35}_{-0.4}$	$1.06^{+0.07}_{-0.10}$	$0.67^{+0.05}_{-0.05}$	$0.13^{+0.21}_{-0.20}$	$0.23^{+0.17}_{-0.17}$	—	$13.74^{+0.43}_{-0.32}$	990^{+270}_{-140}
cgNFW	$3.50^{+0.32}_{-0.36}$	$4.76^{+0.55}_{-0.40}$	412^{+123}_{-143}	$2.39^{+0.39}_{-0.43}$	$7.22^{+0.34}_{-0.40}$	$1.06^{+0.08}_{-0.05}$	$0.67^{+0.03}_{-0.06}$	$0.13^{+0.21}_{-0.20}$	$0.23^{+0.16}_{-0.18}$	$19.00^{+8.38}_{-8.34}$	$13.87^{+0.42}_{-0.35}$	1090^{+220}_{-150}

Table 3.3: Final inference on the parameters for the anisotropic models. We report the maximum-likelihood values of our inferred posterior distributions, along with the 16th and 84th percentiles as a measure of our uncertainty. All quantities are measured in units of solar mass, solar luminosity, kilometers per second and kiloparsecs.

the structure of the dark matter halo and the total stellar mass of the system, with Agnello et al. (2014) finding a stellar mass-to-light ratio $Y_{\star,V} \sim 4.5$ and a super-NFW cusp $\gamma = 1.6$ and Zhu et al. (2014) inferring $Y_{\star,I} = 6.0 \pm 0.3$ (which converts to a V -band measurement of ~ 7.5) under the assumption of a cored halo. There are a number of possible reasons for this: first, while both studies overlapped in the majority of their globular cluster data, each used a different mass model, and it is possible that these may have unnecessarily constrained the inference on the mass. Specifically, Zhu et al. (2014) chose to use a logarithmic potential model for the dark matter, so enforcing a core, whereas Agnello et al. (2014) used a power law for the dark halo, thus requiring a constant slope at all radii. Neither of these allows the halo much flexibility in the central regions. Further, while Agnello et al. (2014) relied solely on the globular clusters, separating them into three independent populations based on their velocities, positions and colours, Zhu et al. (2014) treated all the globular clusters as a single tracer population, but used the same SAURON data as in this study to constrain the stellar mass-to-light ratio. Thus it is possible that Agnello et al. (2014) lacked the data coverage in these central regions that would have permitted a reliable distinction between cusps and cores.

By exploring multiple mass models and data combinations, we are able to reproduce the results of both studies. First, excluding the stellar kinematics and carrying out the inference using just the satellites and globular clusters in a way more similar to that of Agnello et al. (2014), we infer an isotropic gNFW profile with $\gamma = 1.48 \pm 0.22$, thus reproducing their finding of a cusp. The inference is summarised in Figure 3.16. This data combination also provides only a very weak constraint on the stellar mass-to-light ratio, as might be expected given the relatively small number of globular clusters in the centre as compared to the stars. When we compare the predictions of the best-fit model in this case, however, we find that it significantly overpredicts the stellar velocity dispersions in the centre. Adding in the stellar data, then, requires an excavation of the central regions, giving rise to a sub-NFW inner slope more easily reconciled with the assumption of Zhu et al. (2014), although we do still find a lower stellar mass, which could be a consequence of the fact that, unlike Zhu et al. (2014), we include the black hole in our models. The solution of the discrepancy between the two studies is presumably then that it is simply not possible to constrain the whole mass profile using globular cluster data alone. This really shows the importance of combining multiple tracer populations with different characteristic radii.

We also perform the first joint inference on the black hole, stellar and halo mass parameters that has been carried out with the high-resolution NIFS kinematics of the stars in the central $2''$. When these data were presented in Gebhardt et al. (2011), the structure and scale of the halo and stellar distributions were fixed at Murphy et al. (2011)’s values, which in turn were inferred assuming the black hole mass measured in Gebhardt & Thomas (2009), based on lower-resolution data. The black hole mass that we infer is consistent with the Gebhardt et al. (2011) value within uncertainties: for the best cgNFW model, we find $7.22^{+0.34}_{-0.40} \times 10^9 M_{\odot}$.

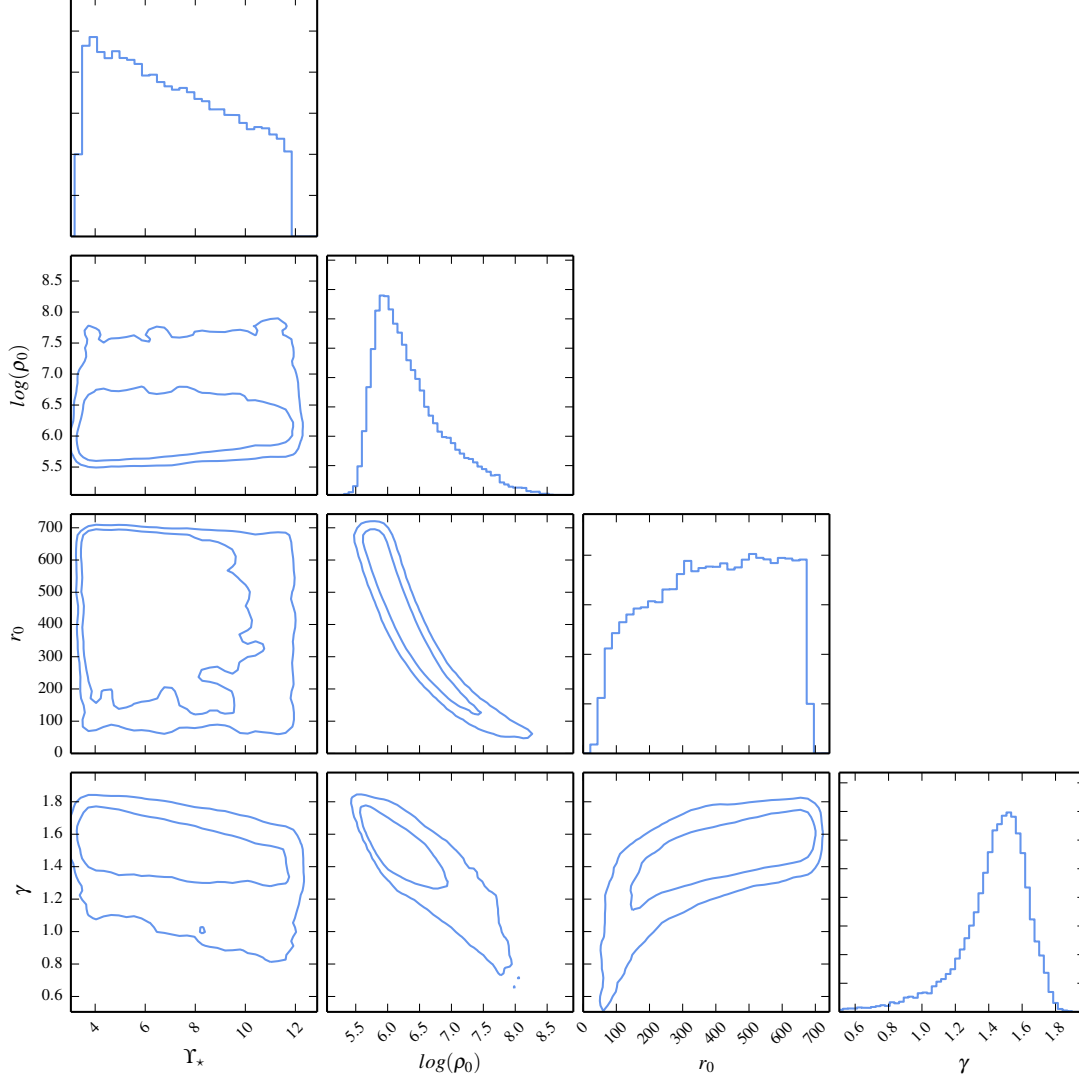


Figure 3.16: Inference on the gNFW model parameters when the stellar data are excluded. The stellar mass-to-light ratio and halo scale radius are unconstrained and the effect of the priors is visible for these parameters, while the halo becomes cuspy, with inner slope $\gamma = 1.48 \pm 0.22$. This is a very different result from the case where all data are modelled in conjunction, and shows the value of the multiple-population method.

3.5.2 The case for a cored halo

It is important to compare our findings for the halo structure with other ETGs; as explained in the Introduction, this is crucial for developing our understanding of the role of baryons in shaping the halo and the diversity of structures that can arise.

A number of other studies have found BCGs to have dark matter profiles which are flatter than NFW: Newman et al. (2011), for instance, combined stellar dynamics with weak and strong gravitational lensing and X-ray data to infer a density slope $\gamma < 1$ for the BCG Abell 383 with 95 % confidence, while Newman et al. (2013) went on to fit gNFW halo models to a sample of seven massive BCGs and found a mean slope of $\gamma = 0.5 \pm 0.1$, inconsistent with the NFW prediction – and with three of those galaxies hitting the $\gamma = 0$ prior. This is also in line with the earlier studies of Sand et al. (2002); Sand et al. (2004); Sand et al. (2008); Newman et al. (2009); more recently, Caminha et al. (2015) have also modelled multiple image families lensed by the cluster Abell S1063 to infer a significant core radius $r_c \sim 100$ kpc. The picture, however, is quite different for elliptical galaxies in the field, and a number of strong lensing studies of ETGs have found halo slopes consistent with NFW or even super-NFW dark matter profiles: Sonnenfeld et al. (2015), for instance, modelled the global properties of a sample of 81 lenses and found the inner slope to be consistent with an NFW profile, while Grillo (2012) performed a similar analysis on a smaller lens sample and found $\gamma = 1.7 \pm 0.5$ ($\gamma = 2 \pm 0.2$) assuming a Salpeter (Chabrier) IMF, which is inconsistent with what we find here (though see also Barnabè et al., 2013, for an example of field ellipticals from the SLACS survey that are found to have sub-NFW inner slopes). At first sight, this might seem to point tentatively to the existence of two very different evolutionary paths for ETGs in these two environments, and the implausibility of a one-size-fits-all dark matter halo profile: however, the sample size of ETGs for which this analysis has been carried out clearly remains too small for any meaningful conclusions on this issue.

We now consider the implications of M87’s central core in the context of the baryonic feedback processes that play a role in the centres of galaxies. Whilst a cored halo is inconsistent with dark-matter-only simulations, there have been a number of recent simulations in which baryonic effects have been included (e.g. Mead et al., 2010; Velliscig et al., 2015; Schaller et al., 2015), and we note some similarities with these. Here, we focus on two of the most recent studies of this type in order to elucidate the physics, though we stress that this comparison is by no means exhaustive and a number of similar studies have been carried out. Firstly, we consider the zoom-in simulations of Laporte & White (2015), which followed two BCGs from redshift $z = 2$ to $z = 0$ and demonstrated that the

effect of repeated dissipationless mergers is to soften the otherwise-NFW-like cusp by $\Delta\gamma \sim 0.3 - 0.4$ on scales of the stellar half-light radius. In this paradigm, infalling satellite galaxies experience dynamical friction from dark matter as they move through the halo, and this transfers energy to the dark matter, causing it to expand and so become less dense in the central regions. While we do not measure a core radius explicitly, in Figure 3.17 we plot the density slope as a function of radius, $d\log\rho/d\log r$, and see that the slope becomes sub-NFW at radii $r < 10$ kpc and that $\gamma \sim 0.6$ (equivalent to $\Delta\gamma \sim 0.4$) at $r \sim 5$ kpc, which is comparable to (though smaller than) the scale radius of the starlight $R_e = 16$ kpc measured by Kormendy et al. (2009). However, we find a cored centre while these simulations find only $0.6 < \gamma < 1$ down to the innermost resolvable radius. Whilst they suggest that the merging of black holes could then give rise to cores in the central 3-4 kpc – with the binary black hole system spiralling inwards and transferring energy to the surrounding matter – their simulations do not include the effects of black holes and so they are unable to test this. It therefore remains unclear as to whether merging events would be capable of totally erasing the cusp.

On the other hand, M87 has a large AGN at its centre and this may also contribute to core formation. The second study that we examine is that of Martizzi et al. (2012), which specifically simulated Virgo-like ETGs using a recipe for AGN feedback. They found cores to develop within the inner 10 kpc. In this scheme, outflows of gas due to the AGN are able to irreversibly modify the gravitational potential of the halo and so cause expansion of both the dark and luminous matter. In that study, they suggest that a combination of AGN feedback and black hole effects are the main contributing processes, though they also note that the large amount of gas expelled by the AGN increases the efficiency of the energy transfer due to dynamical friction. This role for AGN is analogous to the role of supernovae in dwarf spheroids (Pontzen & Governato, 2012). However, while the core radius of 10 kpc seems consistent with the density profile that we infer, Martizzi et al. (2012) predict cores in the 3D stellar density, whereas M87 is known to have a ‘cuspy core’ in its 2D surface brightness ($\gamma_{2D} = 0.18$), making it considerably cuspy in 3D. Indeed, the finding of cores in the 3D density distribution of the stars seems at odds with observations of large numbers of ETGs (e.g. Kormendy et al., 2009). Again, this seems to be a sign of the complex nature of the processes governing the shape of the density profile, and the need for further investigations to be carried out.

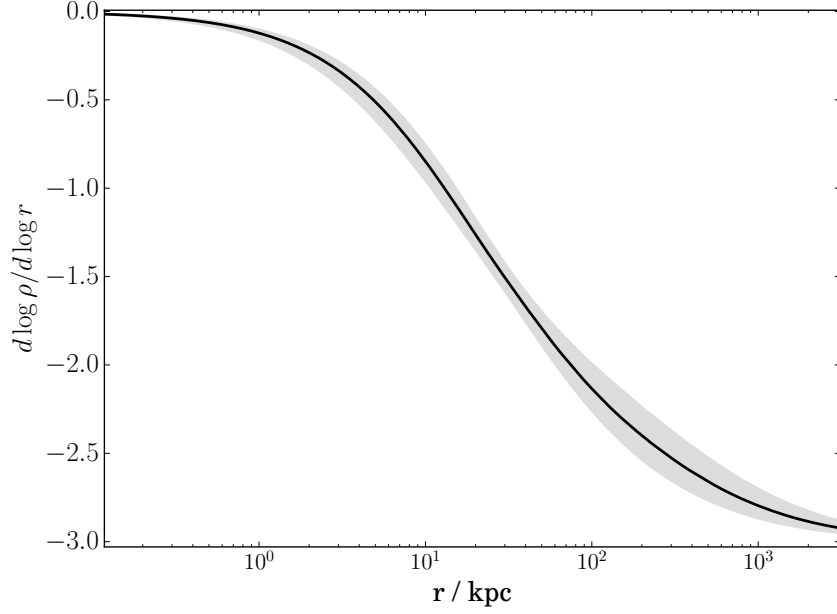


Figure 3.17: The slope of the halo density profile as a function of radius for the anisotropic cgNFW model. The halo becomes sub-NFW ($\gamma < 1$) at $r \sim 10$ kpc and continues to flatten quite rapidly, with $\gamma \sim 0.6$ by $r \sim 5$ kpc.

3.5.3 Implications for the stellar initial mass function

One of the initial motivations for this investigation was to resolve the striking discrepancy between the $Y_{\star,V}$ inferred by Agnello et al. (2014) and Zhu et al. (2014), particularly in light of recent evidence that massive ETGs may have Salpeter-like stellar mass-to-light ratios (e.g., van Dokkum & Conroy, 2010; Auger et al., 2010b; Cappellari et al., 2012). In our isotropic cgNFW model, we find $Y_{\star,V} = 6.9 \pm 0.1$, whereas our anisotropic model finds a significantly lower $Y_{\star,V} = 3.50^{+0.32}_{-0.36}$; on the other hand, previous studies have inferred values even higher than our isotropic model: Zhu et al. (2014) find $Y_{\star,V} \sim 7.5$, and Murphy et al. (2011) finds $Y_{\star,V} \sim 9.9$ when converted to our assumed distance. As noted in Section 3.5.1, ours is the first study to *simultaneously* infer the mass properties of the black hole, stars and dark halo, and from this perspective should lead to smaller degeneracies between the different components; on the other hand, the studies of Zhu et al. (2014) and Murphy et al. (2011) used orbit-based modelling rather than a Jeans analysis, which makes their models able to reflect more flexible and realistic anisotropy structures. It is therefore possible that the large discrepancy between our anisotropic model and previous studies is due to the limitations of the anisotropy profiles that we have assumed; since orbit-based studies have tended to find M87 to have a largely isotropic stellar population, this would

then also explain why our isotropic result is more in line with their results. Indeed, we note that the total mass-to-light ratio – that is, including both dark and light mass – of our isotropic model is broadly consistent with previous results, including Cappellari et al., 2006, whereas the total-mass-to-light ratio in our anisotropic model is not. In any case, we can use the single stellar population models from Bruzual & Charlot (2003, BC03) to understand the implications of the stellar mass-to-light ratio for the age and IMF of M87.

We measure colours for M87 from the CFHT photometry and compare these with the BC03 models. In particular, we determine the mean surface brightness within a $2''$ annulus at $30''$ (i.e., ~ 2.5 kpc) from the centre of M87, masking globular clusters and other artefacts, for each of the *ugriz* filters. We find extinction-corrected $u - g$, $g - r$, $g - i$, and $g - z$ colours of 1.48, 0.67, 1.13, and 1.45, respectively, which agree well with the colours calculated over a much larger aperture from SDSS photometry (Brown et al., 2014). Then, assuming a Salpeter IMF and solar metallicity, the BC03 models give an age of 10 Gyr and $Y_{*,V} \approx 5.7$, significantly lower than the Zhu et al. (2014) or Murphy et al. (2011) values and thus implying the need for an IMF even more extreme than that found by Conroy & van Dokkum (2012b) if these high mass-to-light ratios are to be explained. If we instead assume a Chabrier IMF and solar metallicity, the BC03 models still find a ~ 10 Gyr age, but $Y_{*,V} \approx 3.7$, which is in good agreement with our dynamical measurement when we allow for anisotropy. We note, however, that our dynamical inference on $Y_{V,*}$ is very sensitive to our assumptions about anisotropy. Furthermore, our assumption of solar metallicity – and, indeed, a constant $Y_{*,V}$ with radius – may be inappropriate. For example, Montes et al. (2014) use UV to NIR photometry to demonstrate that M87 has colour gradients, and they interpret these as being the result of a metallicity gradient.

3.5.4 Importance of correctly modelling the underlying tracer populations

The Jeans equation requires us to know the underlying tracer density for each population, and this, along with the mass profile and anisotropy, determines the velocity dispersion that is measured. Therefore any uncertainty or bias in the parameterisation of the tracer density might manifest itself as uncertainty or bias in the resulting velocity dispersions and, in a study such as this in which the velocity dispersions are the primary tool for the inference, also as uncertainty or bias in the mass profile. A particular worry here is the distributions of the globular clusters, as the spectroscopic catalogue of Strader et al. (2011) is only a subsample of the wider population and, by virtue of its being selected for spectroscopy, is subject to some non-trivial selection function that may change the apparent

distribution from that of the true underlying one. It is therefore important to model the distribution using an independent photometric sample. Indeed, this was the motivation for the initial photometric study of Chapter 2.

To demonstrate the impact on the inference of using the density of the spectroscopic subsample in the Jeans analysis rather than that of the more representative photometric sample, we repeat our analysis using the three-population decomposition given in Agnello et al. (2014). This decomposition was carried out based solely on the Strader et al. (2011) spectroscopic catalogue, and separates the globular clusters into a compact red component along with two much more extended blue and intermediate-colour components. Our analysis differs from Agnello et al. (2014)’s in that they focussed mainly on virial decompositions and used the globular cluster data exclusively, whereas we use a spherical Jeans analysis and combine the globular clusters with the stellar kinematics and satellite galaxy datasets. When we carry out the inference using the three populations, we find that we underpredict the halo mass in the region in which the globular clusters dominate the fit by ~ 0.25 dex. M_{vir} also decreases by $\sim 10\%$, though the effect is presumably smaller here because the globular cluster data only extend out to a fraction of the virial radius. Unsurprisingly, our inference on the stellar mass-to-light ratio is robust against the globular cluster Sérsic distributions, as this is mainly determined by the properties of the stellar population, which is fixed by our fit to the surface brightness profile.

Another possible bias in spherical Jeans modelling could result from any flattening in the potential: indeed, there have been recent suggestions that the shape of a galaxy’s dark matter halo may correlate with the shape of its luminous matter (e.g. Wu et al., 2014), and, while the modelling of Chapter 2 assumed the globular clusters to be spherically distributed, they do exhibit some ellipticity at larger radii. However, the effect of this was analysed in Zhu et al. (2014) using a series of axisymmetric Jeans models, and it was found that the impact of any realistic amount of flattening on the mass profile was small, with up to a 10 % decrease in the mass for globular clusters with elliptical profiles. This is therefore unlikely to be a major concern here.

3.5.5 Anisotropy, and a pinch radius from multiple tracer populations

As explained in Section 4.3, we incorporated anisotropy in our models by assuming a scaled Osipkov-Merritt profile for the stars and constant anisotropies for the globular cluster populations, and inferred mild radial anisotropies for the three populations, which in turn led to changes in our inference on the mass model relative to the isotropic case. As the inclusion of the anisotropy parameters is considerably more expensive from a

computational point of view, it is of interest to investigate ways of obtaining robust mass estimates using models with fixed anisotropy. It is well known (e.g. Wolf et al., 2010) that, for the Jeans analysis of a single tracer population, there exists a pinch radius at which the dependence of the inferred mass profile on the anisotropy is minimal. This occurs at approximately the 3D effective radius of the tracer and can be quoted as a robust measure of the enclosed mass – instead of an entire radial profile – when it is not possible to infer the anisotropy from the data. However, our case is slightly more complicated than those studied by the previous authors as we are combining three completely independent tracer populations. It is therefore of great interest to test whether the pinch radius is still recovered in this case.

For simplicity, we run a series of models in which the stellar anisotropy is fixed at some non-zero constant value while the globular clusters remain isotropic. This is guided by our earlier results, which found the stars to be considerably more anisotropic than the globular clusters. To span a range of anisotropies, we let the anisotropy run from $\beta = -1$ (corresponding to the case where the velocity dispersion in the radial direction is twice that in the tangential direction) to $\beta = 0.5$ (at which point the tangential velocity dispersion is half the radial velocity dispersion). The results are shown in Figure 3.18. Indeed, we do find the three β -curves to intersect at a single radius, as in the case of single-tracer population models: however, this occurs at a radius $= 111 \pm 4$ kpc (with mass $\log(M(r < r_p)/M_\odot) = 13.09 \pm 0.40$), which is much too large to be associated with the effective radius of the starlight. This is presumably a consequence of our use of additional tracer populations, all at larger radii than the stars. This is an interesting finding which will be useful for further multiple-population studies of this kind.

3.6 Conclusions

We have modelled the mass profile of M87 using a combination of stellar, globular cluster and satellite galaxy kinematics in a Jeans analysis, and our main conclusions are as follows:

1. M87 is a massive BCG with a sub-NFW dark matter halo. Under assumptions of isotropy, the scatter introduced by the use of different models to parameterise the halo is small. In this paradigm, we find M87 to have a high stellar mass-to-light ratio of $Y_\star = 6.9 \pm 0.1$ in the V -band and a cored dark matter halo, with inner slope $\gamma < 0.14$ at the 95 % confidence level.

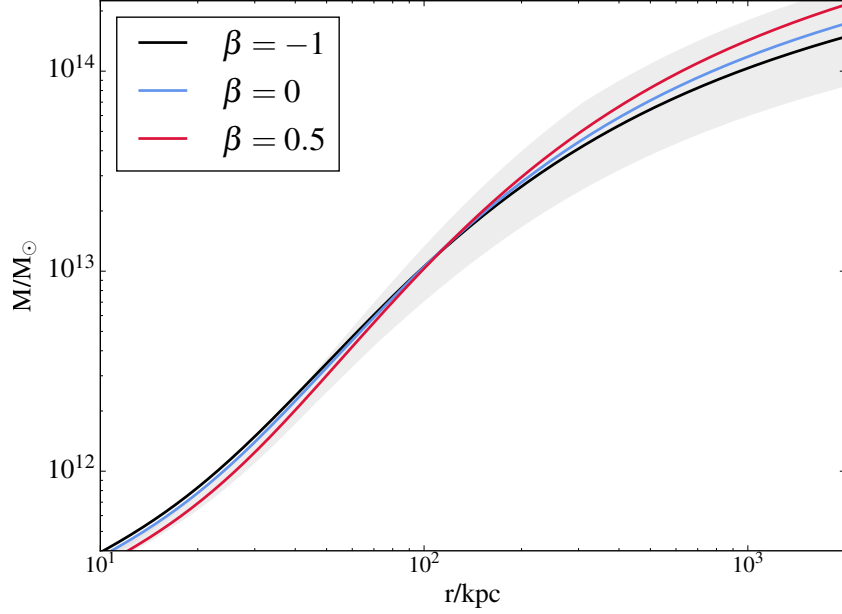


Figure 3.18: Varying the stellar anisotropy has an effect on the inferred mass profile at all radii except for the pinch radius, clearly visible at ~ 111 kpc, where the dependence of the mass on the anisotropy is minimised. While such an extreme tangential bias as $\beta = -1$ is not physically motivated, we run the inference for a range of constant stellar anisotropies between $\beta = -1$ and $\beta = 0.5$ to demonstrate the emergence of the pinch-point.

2. When the model is modified to allow each tracer population some element of anisotropy, differences arise between the halo models, and a cored generalised NFW profile provides the best description of the dynamics. This reduces the stellar mass-to-light ratio to $Y_{\star} = 3.50^{+0.55}_{-0.36}$ and slightly relaxes the constraint on the inner slope, with a core radius $r_c = 19.0 \pm 8.3$ kpc. All three tracer populations are characterised by mildly radially anisotropic orbits.
3. The V-band stellar mass-to-light ratio of M87 is consistent with a picture in which its stellar populations are old (~ 10 Gyrs), with solar metallicity. In the isotropic case, this implies a Salpeter-like IMF, whereas when anisotropy is accounted for, a Chabrier-like IMF is preferred.
4. The inclusion of tracers at a variety of spatial scales has a significant impact on the inference. Modelling the mass using the globular clusters and satellite galaxies alone - that is, without the stars - we infer a cuspy halo, with inner slope $\gamma \sim 1.5$, but when the stars are included in the inference, the halo becomes sub-NFW. This shows the importance of consistently modelling the profile across a range of spatial scales.

5. It is important to properly characterise the distributions of the underlying tracer populations as opposed to those of their kinematic subsamples. We have shown that the use of globular cluster colour and spatial distributions based on the kinematic dataset alone leads to a systematic underprediction of the total mass of the system.
6. When the inference is carried out with different (constant) values for the stellar anisotropy between -1 and 0.5, a pinch radius clearly emerges at which the dependence of the enclosed mass on the anisotropy is minimised. This gives an estimate of the enclosed mass $\log(M(r < 110\text{kpc})/M_{\odot}) = 13.09 \pm 0.40$ which should be robust against assumptions about anisotropy.

THE MOST MASSIVE GALAXIES IN CLUSTERS ARE ALREADY FULLY GROWN AT $z \sim 0.5$

Abstract

By constructing scaling relations for galaxies in the massive cluster MACSJ0717.5 at $z = 0.545$ and comparing with those of Coma, we model the luminosity evolution of the stellar populations and the structural evolution of the galaxies. We calculate magnitudes, surface brightnesses and effective radii using HST/ACS images and velocity dispersions using Gemini/GMOS spectra, and present a catalogue of our measurements for 17 galaxies. We also generate a photometric catalogue for ~ 3000 galaxies from the HST imaging. With these, we construct the colour-magnitude relation, the fundamental plane, the mass-to-light versus mass relation, the mass-size relation and the mass-velocity-dispersion relation for both clusters. We present a new, coherent way of modelling these scaling relations simultaneously using a simple physical model in order to infer the evolution in luminosity, size and velocity dispersion as a function of redshift, and show that the data can be fully accounted for with this model. We find that (a) the evolution in size and velocity dispersion undergone by these galaxies between $z \sim 0.5$ and $z \sim 0$ is mild, with $R_e(z) \sim (1+z)^{-0.40 \pm 0.32}$ and $\sigma(z) \sim (1+z)^{0.09 \pm 0.27}$, and (b) the stellar populations are old, ~ 10 Gyr, with a ~ 3 Gyr dispersion in age, and consistent with evolving purely passively since $z \sim 0.5$ with $\Delta \log M/L_B = -0.55^{+0.15}_{-0.07} z$. The implication is that these galaxies formed their stars early and subsequently grew dissipationlessly so as to have their mass already in place by $z \sim 0.5$, and suggests a dominant role for dry mergers, which may have accelerated the growth in these high-density cluster environments.

4.1 Introduction

ETGs obey tight scaling relations, such as the fundamental plane, out to redshifts $z \sim 1$ (Djorgovski & Davis, 1987; Dressler et al., 1987; see also e.g. Holden et al., 2010), with a small amount of scatter which suggests that their stellar populations formed early, then faded passively. On the other hand, ETGs at $z \sim 2$ are observed to be extremely compact (e.g. van Dokkum et al., 2008), and must therefore evolve through a series of minor mergers and accretion events into the large, passive systems that we see today. Reconciling these constraints on their structural and stellar evolution remains a key challenge, and, in particular, requires that well-established scaling relations are re-examined in new detail.

The fundamental plane – which can be accounted for by the assumption that ETGs are virialised, near-homologous systems – relates the effective radius R_e , velocity dispersion σ and surface brightness $\langle \mu_e \rangle$ of ETGs as

$$(4.1) \quad \log R_e = \alpha \log \sigma + \beta \langle \mu_e \rangle + \gamma,$$

with $\alpha = 1.24$ and $\beta = 0.33$ in the local Universe (Jørgensen et al., 1995). Whilst the tilt of the fundamental plane has been shown to remain constant with redshift, the evolution of the zeropoint has been taken, in the past, as evidence for the passive fading of the stellar populations and used to measure stellar age (e.g. van Dokkum & Franx, 1996; Bender et al., 1996; van Dokkum & Ellis, 2003; Jørgensen et al., 2006), generally implying a mean star formation redshift $z_f > 2$. However, the fundamental plane is also sensitive to evolution in velocity dispersion and size, and whilst more recent studies have attempted to account for this, no consensus has yet been reached on the strength of this evolution. For instance, Saglia et al. (2010) measured the structural evolution of a sample of field galaxies out to $z \sim 0.9$ and found it to be significant, with its inclusion in the zeropoint analysis increasing the stellar age by 1-4 Gyr (depending on morphology and redshift). On the other hand, Jørgensen & Chiboucas (2013) and Jørgensen et al. (2014) found only very small differences in size and velocity dispersion for cluster galaxies across similar redshifts – roughly one-third that of Saglia et al. (2010) – and made the suggestion that structural evolution may depend on environment, with accelerated growth in dense clusters. However, the comparison of field and cluster galaxies of Newman et al. (2014) found no evidence for an environmental dependence of their sizes at $z = 1.8$, indicating that any changes must imprint themselves in a narrow redshift window. The question of how and when any growth occurs, then, and how it relates to the stellar populations, remains open, and it is important to try to answer this further using independent samples and methods.

Another extremely simple but useful scaling relation that can be used to give a measure

of the ETG formation epoch is the colour–magnitude relation (CMR; Baum, 1959; Sandage, 1972; Visvanathan & Sandage, 1977), in which the ETGs fall along a tight red sequence, whose intrinsic scatter is mainly determined by the distribution of stellar age. In contrast to the fundamental plane, which contains information on the galaxy mass structure, the CMR depends almost wholly on the properties of the stellar populations. Several studies out to redshifts $z \sim 1.3$ have shown the intrinsic scatter about the CMR to be consistently small (generally < 0.1 mag: see Stanford et al., 1995; Ellis et al., 1997; Bower et al., 1998; Stanford et al., 1998; Mei et al., 2009), suggestive of generally old stellar populations, with a redshift of formation $z_f > 2$ and a small spread in age. However, it is hard to disentangle these two degenerate factors, given that the scatter decreases with both increasing age – as stars become asymptotically redder – and increasing synchronicity – as stars with similar ages have similar colours. It is therefore possible for a stellar population with recent, synchronised star formation to have the same small scatter as one in which the star formation happened longer ago but was dispersed. Clearly, if this degeneracy can be broken, it can provide informative complementary constraints on the fundamental plane.

The aim of this chapter is to construct these scaling relations for the cluster MACSJ0717.5+3745 (hereafter MACSJ0717) at $z=0.545$, and to compare with those of the Coma galaxies to investigate their evolution in terms of galaxy structure and stellar populations. The chapter is organised as follows: in Section 4.2 we introduce the data and explain our reduction methods; in Section 4.3 we construct the scaling relations and in Section 4.4 we interpret these in terms of stellar population models. Section 4.5 and Section 4.6 then give a discussion and summary.

4.2 Data sources and reduction

4.2.1 Sources

For both MACSJ0717 and Coma, we construct the colour-magnitude relation (CMR), the fundamental plane and mass-to-light versus mass (MLM) relation and the mass-effective-radius (MR) and mass-velocity dispersion (MS) relations. We therefore require colours, surface photometry and kinematics in each case.

For MACSJ0717, we measure photometry using archival HST/ACS images that were observed as part of the Cluster Lensing And Supernova survey with Hubble (CLASH; Postman et al., 2012). For size and surface brightness measurements, we use the F475W (exposure time: 4064s) and F625W bands (exposure time: 4128s) in order to bracket the Balmer break in the rest frame. For catalogue generation using SExtractor (Bertin,

1996), we additionally use the multiband image (summing all 16 CLASH filters) as the reference image for object detection (as explained in more detail in Section 4.2.2). These data are available from the CLASH archive, and have been previously corrected for galactic extinction and redrizzled to a pixel scale of $0.065''/\text{pixel}$. We measure velocity dispersions using Gemini/GMOS spectra, available in the Gemini archive for a subsample of 31 galaxies. These spectra, along with associated flat and bias frames, were taken over four dates between 04/02/2003 and 02/03/2003, using GMOS in multi-object mode with the B600_G5303 grating (which has a resolution $R = 1688$ for a $0.5''$ slit width at 461nm), as part of the science program GN-2002B-Q-44.

For Coma, we use the kinematic data from Jørgensen et al. (1999a), integrated photometry, observed in the Johnson U and V bands, from Terlevich et al. (2001), and the surface photometry in the Gunn r' -band from Jørgensen et al. (1995). We refer the reader to these papers for further information, though we note that the photometry has been previously corrected for galactic extinction, and that we convert the Gunn r' photometry to AB magnitudes using the corrections listed in Frei & Gunn (1994).

4.2.2 Photometry

The CLASH database (Postman et al., 2012) provides a SExtractor-generated catalogue of isophotal magnitudes for ~ 8000 objects detected in the HST images; however, in order to calculate colours precisely for the CMR and avoid under-estimating the flux of the largest low-surface-brightness galaxies, we use SExtractor to generate a catalogue of integrated aperture magnitudes, using the multiband image for source detection and the single-band images for measurement. We use an aperture radius of $1.3''$ in both wavebands, and select objects according to the following criteria:

1. $\text{STELLARITY} < 0.95$ & $\text{FWHM} > 0.2''$, to avoid contamination from stars;
2. $\Delta F_{475W} < 0.2$ mag, $\Delta F_{625W} < 0.2$ mag & $\text{FLAG} > 4$, to avoid objects that may not have been properly deblended or do not have reliable photometry;
3. $F_{625W} < 25$ mag, as a luminosity cut-off.

We cross-correlate our SExtractor catalogue with the existing CLASH catalogue, which provides photometric redshifts based on the full set of 16 ACS/HST filters, and further reject all objects whose redshift range do not satisfy $z_{\min} < 0.545 < z_{\max}$ for the 95% confidence intervals z_{\min} and z_{\max} .

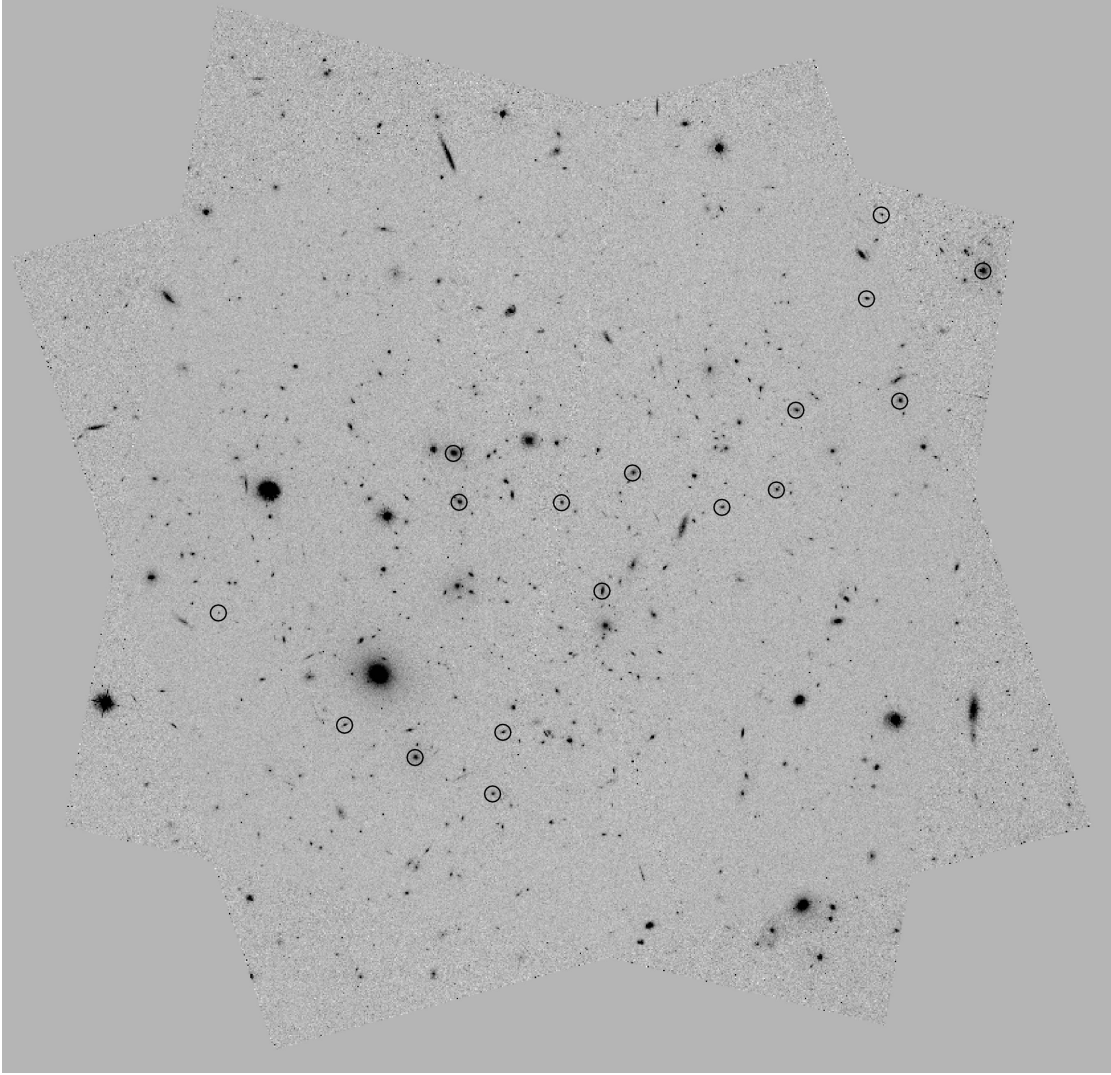


Figure 4.1: HST/ACS F625W image of MACSJ0717, with a pixel scale of $0.065''/\text{pixel}$ and dimensions $325 \times 325''$; black circles show the final sample of 17 galaxies.

We model the surface brightnesses of the galaxies in the F625W image using single-component Sérsic profiles, with

$$(4.2) \quad I(R) = I_e \exp \left[-k_n \left(\frac{R}{R_e} \right)^{\frac{1}{n}} - 1 \right]$$

for Sérsic index n , effective radius R_e , $k_n = 2n - 0.324$ and radius $R^2 = q^2(x - x_0)^2 + (y - y_0)^2/q^2$ for axis ratio q and galaxy centroid (x_0, y_0) (i.e. both R_e and R are circularised, projected quantities). To test for and eliminate systematics in our size measurements, we proceed via two different methods. First, we use the curve-of-growth fitting (COG) code presented in Houghton et al. (2012), which has been rigorously tested and shown to reproduce simulated images to high accuracy (see the appendix of that paper). This masks bright regions close to the galaxy to create a ‘clean’ cutout, whose circularised integrated light profile is then fitted using a chi-squared minimisation. Second, we use a surface-fitting code based on that developed for Chapter 5, which explores the six-dimensional parameter space represented by $(x_0, y_0, R_e, n, q, \theta)$ – where (x_0, y_0) is the centroid of the Sérsic profile, θ is the position angle and q the ellipticity – using the MCMC sampler EMCEE (Foreman-Mackey et al., 2013). In this case, for objects with crowded fields we model all bright objects simultaneously. For both the COG and the surface fitting routines, we convolve the model with the radial profile of an unsaturated star in the image to account for the PSF, and we find very good agreement between the two sets of results. For Coma, we use the sizes and surface brightnesses presented in Jørgensen et al. (1995), which were measured using de Vaucouleurs COGs.

Of the 31 MACSJ0717 objects in the spectroscopic sample, only 19 of these overlap with the CLASH field; of these, a further two are in fact stars (as can be clearly seen in both the imaging and the spectra), and so are excluded from the analysis. We therefore end up with a final sample of 17 galaxies. The completeness of the spectroscopic sample relative to the galaxy population is illustrated in Figure 4.2. Three of the systems in the final sample (object IDs 11, 13 and 22) have clear extended stellar haloes or bulge+disk morphologies which mean that we are unable to construct satisfactory models using a single Sérsic component; for these, we add a second component in the surface-fitting method and find that this enables us to model their light profiles down to the noise. It is important to do this sparingly in order that all our measured sizes are directly comparable.

The galaxy cutouts and model residuals are presented in Figure B.1 in the Appendix, and the effective radii and surface brightnesses are included in Table 4.1, with the latter corrected for cosmological dimming and extinction. Figure 4.1 shows the CLASH footprint, with our final galaxy sample marked.

4.2.3 Spectroscopy

We reduce the spectra of all 31 galaxies using the GMOS package in IRAF (Tody, 1993), calibrating the wavelength using the skylines in the exposures, according to the UVES sky emission atlas (Hanuschik, 2003). To attain a higher signal-to-noise, we stack the 12 exposures slitlet by slitlet.

We model each spectrum as the sum of a galaxy and a continuum component. For the former, we use stellar templates for G, K, A and F stars from the Indo-US Stellar Library of Coudé Feed Stellar Spectra (Valdes et al., 2004), which we redshift and convolve with a dispersion $\sigma_{model}^2 = \sigma_{true}^2 + \sigma_{inst}^2 - \sigma_{tmp}^2$ where σ_{true} is the physical velocity dispersion of the system, σ_{inst} is the instrument resolution and σ_{tmp} is the intrinsic resolution of the templates (which is 1.2 \AA for the Indo-US templates). We measure the resolution in each spectrum by fitting Gaussians to the skylines, and find this to be constant across the slitlets, with $\lambda/d\lambda = 3030$. The continuum is an additive order-6 polynomial which accounts for the difference in shape between the templates and the true spectrum, and regions where atmospheric absorption dominates the spectrum are masked. We therefore have two free non-linear parameters – the redshift and velocity dispersion for the galaxy – and 15 linear parameters – the weights of each of the nine stellar templates, and the coefficients of the order-6 polynomial. We construct a likelihood of the data given the model

$$(4.3) \quad \ln L = -\frac{1}{2} \sum_k \left(\frac{F_{k,obs} - F_{k,mod}}{\delta F_k} \right)^2$$

where the sum is over k pixels along the wavelength axis, and F_{mod} , F_{obs} and δF are the model flux, observed flux and observed flux uncertainty respectively. We then explore the posterior probability distribution of the model given the data using MCMC sampling.

Our kinematic models are shown in the far-right panels of Figure B1 and the resulting velocity dispersions are included in Table 4.1. We test the robustness of our kinematic inference by repeating the exercise using the lower-resolution galaxy templates of Bruzual & Charlot (2003) and find that the typical uncertainty in the velocity dispersion is of order 5%; as this is larger than our statistical uncertainties, we impose this as the uncertainty on all our velocity dispersion measurements (though only the statistical uncertainties are given in Table 4.1). We also check the robustness of our method by modelling the spectra independently using the penalised pixel fitting (pPFX) software of Cappellari & Emsellem (2004, v4.15, also using the INDO-US library), and find the uncertainty to be less than 5%. To our inferred velocity dispersions we apply aperture corrections following the prescription in Jørgensen et al. (1995), correcting the dispersions from the $1''$ apertures

ID	RA/deg	DEC/deg	R_e/kpc	σ/kms^{-1}	$\langle\mu_e\rangle/\text{mag}$	F625W/mag	F475W - F625W /mag	$\log(M_\star/M_\odot)$	$\log(M_{dyn}/M_\odot)$
2	109.4079	37.7363	3.04 ± 0.18	220.9 ± 4.0	19.92 ± 0.11	21.46 ± 0.03	1.88 ± 0.03	11.40 ± 0.06	11.24 ± 0.03
4	109.3952	37.7316	3.42 ± 0.20	198.8 ± 5.4	19.93 ± 0.10	21.24 ± 0.04	1.86 ± 0.03	11.42 ± 0.09	11.20 ± 0.03
5	109.3943	37.7358	1.69 ± 0.06	220.2 ± 4.8	18.44 ± 0.05	21.27 ± 0.01	1.80 ± 0.02	11.28 ± 0.09	10.98 ± 0.01
7	109.3981	37.7515	4.06 ± 0.12	298.7 ± 3.0	19.57 ± 0.05	20.49 ± 0.05	1.89 ± 0.02	11.73 ± 0.11	11.63 ± 0.02
8	109.3986	37.7548	9.51 ± 0.30	157.8 ± 2.4	20.68 ± 0.54	18.48 ± 0.01	1.90 ± 0.01	12.51 ± 0.11	11.44 ± 0.02
9	109.3857	37.7454	14.55 ± 0.33	274.7 ± 4.7	21.37 ± 0.02	19.54 ± 0.01	1.70 ± 0.01	11.61 ± 0.15	12.11 ± 0.01
10	109.3892	37.7514	2.68 ± 0.11	159.1 ± 3.6	19.25 ± 0.07	21.07 ± 0.02	1.75 ± 0.03	11.22 ± 0.10	10.90 ± 0.02
11	109.3831	37.7535	2.52 ± 0.97	184.8 ± 3.9	20.51 ± 0.09	20.59 ± 0.03	1.77 ± 0.02	11.42 ± 0.10	11.00 ± 0.03
12	109.3707	37.7523	5.07 ± 0.24	216.5 ± 4.1	20.44 ± 0.08	20.88 ± 0.02	1.79 ± 0.02	11.39 ± 0.09	11.44 ± 0.02
13	109.3690	37.7577	2.41 ± 0.88	109.4 ± 3.1	20.66 ± 0.08	20.84 ± 0.05	1.69 ± 0.03	11.08 ± 0.15	10.52 ± 0.03
14	109.3601	37.7584	10.26 ± 0.94	223.2 ± 4.1	21.28 ± 0.16	20.20 ± 0.05	1.85 ± 0.02	11.80 ± 0.11	11.77 ± 0.04
16	109.3529	37.7672	11.31 ± 0.42	302.8 ± 3.8	20.82 ± 0.05	19.52 ± 0.04	1.89 ± 0.02	11.99 ± 0.12	12.08 ± 0.02
22	109.4019	37.7341	3.40 ± 0.38	205.1 ± 5.9	20.79 ± 0.02	20.22 ± 0.01	1.97 ± 0.02	12.01 ± 0.08	11.22 ± 0.02
23	109.3754	37.7511	3.33 ± 0.13	227.1 ± 6.7	19.88 ± 0.05	21.25 ± 0.02	1.92 ± 0.03	11.57 ± 0.07	11.30 ± 0.01
24	109.3629	37.7653	3.79 ± 0.17	187.6 ± 4.7	20.14 ± 0.08	21.21 ± 0.02	1.80 ± 0.03	11.31 ± 0.10	11.19 ± 0.02
28	109.4188	37.7439	0.21 ± 0.00	166.1 ± 5.2	17.52 ± 0.18	24.77 ± 0.14	1.34 ± 0.03	8.51 ± 0.06	9.84 ± 0.01

Table 4.1: Photometry, kinematics and stellar and dynamical masses for the galaxies in the final MACSJ0717 sample. The photometry has been corrected for extinction and cosmological dimming; $F625W$ magnitudes are calculated from surface fitting the HST/ACS image while $F475W - F625W$ colours are calculated from the SExtractor catalogues, as described in Section 4.2.2. We discuss the fact that some objects have $M_\star > M_{dyn}$ in Section 4.5.3.

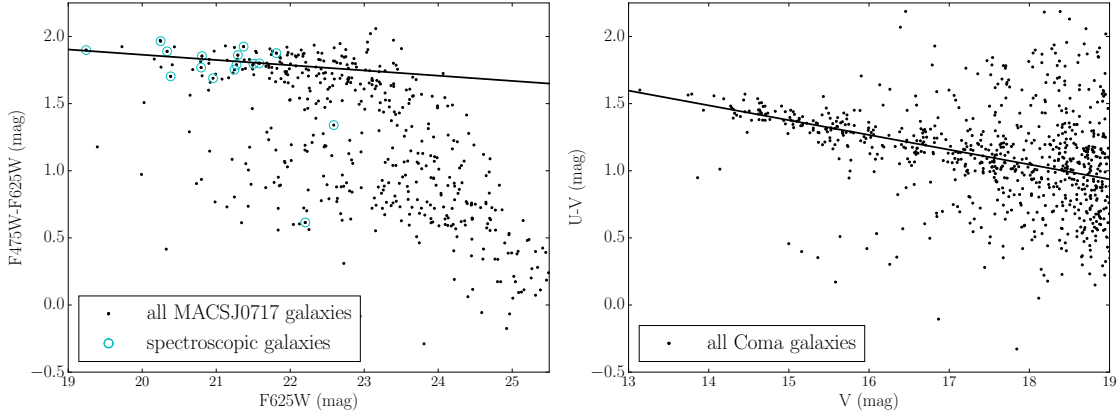


Figure 4.2: The colour-magnitude relations for MACSJ0717 (left) and Coma (right), with all apparent magnitudes measured in the observed frame. The 17 galaxies included in the MACSJ0717 spectroscopic sample are circled in turquoise, and indicate that our spectroscopic subsample is 50% complete down to an absolute magnitude of -21.5 mag in the F625W band. The fits to the red sequences are also plotted.

over which we extracted the spectra to a standard aperture size of $3.4''$ at the distance of Coma.

4.3 Scaling relations

4.3.1 The Colour-Magnitude Relation

We use a mixture model (Hogg et al., 2010) to fit for the slope α_{CMD} , intercept β_{CMD} and intrinsic scatter σ_{CMR} of the red sequence according to the equation

$$(4.4) \quad M_1 - M_2 = \alpha_{CMD} M_2 + \beta_{CMD},$$

where M_1 and M_2 represent the ‘blue’ and ‘red’ magnitudes respectively, with $M_1, M_2 = U, V$ for Coma and $M_1, M_2 = F475W, F625W$ for MACSJ0717. Our model assigns every point a probability of belonging to either the linear distribution of the red sequence or a distribution of outliers that is Gaussian in colour, and seeks the best model parameters via an MCMC exploration.

We test this routine by applying it to the data for Coma, and obtain a value for the scatter $\sigma = 0.065 \pm 0.009$, which agrees with the value $\sigma = 0.069 \pm 0.01$ quoted in Terlevich et al. (2001). We note that, in that study, different morphological groupings of galaxies have their scatters measured separately and give significantly different results: however, in our sample we do not make any cuts based on morphology. Also, we find that our model is very sensitive to the choice of upper magnitude limit, with stricter magnitude limits leading to smaller inferred scatters. This may be a sign of the inadequacy of a single Gaussian for describing the outlier distribution. To compare the scatter of the two clusters, then, it is

important to cut both samples at the same physical magnitude, and it is the smaller sample of higher-redshift galaxies, with generally fewer galaxies at the faint end, that dictates where this should be. This sets an apparent magnitude cut-off for the higher-redshift cluster of $F625W = 24$ mag, which we convert to an r -band cut-off for the Coma galaxies using the stellar population models of Section 4.4.

We find the red sequence of MACSJ0717 to have a shallower slope than Coma, consistent with the fact that Coma is being observed in bluer filters. The scatter is consistent with that of Coma, though slightly smaller, indicating that the stellar populations are already old and red in the MACSJ0717 galaxies, as explored in Section 4.4.1 and discussed in Section 4.5.1. This may also be affected by the fact that a smaller interval sampled by the U and V filters in Coma's rest frame than by the MACSJ0717 filters. We also note that the MACSJ0717 galaxies appear to extend to brighter magnitudes than the Coma sample (the brightest Coma galaxy has an absolute magnitude $M_U = -20.2$, whereas the MACSJ0717 galaxies extend to $M_U = -19.42 \pm 0.18$); though this is partly an effect of the different filters with which each cluster has been observed, the filters and redshifts are such that this discrepancy would in fact be *amplified* by a filter correction. This is a point that we return to in the analysis of other scaling relations in Section 4.3.3. The fitted red sequence of MACSJ0717 is shown in Figure 4.2 and a summary of the results is given in Table 4.1.

4.3.2 The Fundamental Plane

Initially, we fit the fundamental planes for the two clusters independently; as we find their slopes to be consistent with no evolution, we then model them simultaneously, requiring both to be parallel, in order to infer the offset between the two. We assume each dataset to lie on a plane

$$(4.5) \quad \begin{aligned} \log R_{e,z} &= \alpha_{FP,z} \log \sigma_z + \beta_{FP,z} \langle \mu_{e,z} \rangle + \gamma_{FP,z} \\ \log R_{e,0} &= \alpha_{FP,0} \log \sigma_0 + \beta_{FP,0} \langle \mu_{e,0} \rangle + \gamma_{FP,0} \end{aligned}$$

where the subscript x_0 refers to quantities relating to the low-redshift cluster (Coma), and the subscript x_z corresponds to the high-redshift cluster. To reduce the scale of the degeneracy between the plane parameters, we redefine $\sigma_{FP} = \sigma/100 \text{ km s}^{-1}$ and $\langle \mu_{e,FP} \rangle = \langle \mu_e \rangle - 20$. We further assume the independent variables to be drawn from a multivariate Gaussian distribution $N(\tilde{v}_{FP}, \tilde{\tau}_{FP})$ with mean $\tilde{v}_{FP} = (v_{\log \sigma}, v_{FP, \mu_e})$ and variance

$$(4.6) \quad \tilde{\tau}_{FP}^2 = \begin{pmatrix} \tau_{\log \sigma}^2 & \rho \tau_{\log \sigma} \tau_{\mu_e} \\ \rho \tau_{\log \sigma} \tau_{\mu_e} & \tau_{\mu_e}^2 \end{pmatrix}.$$

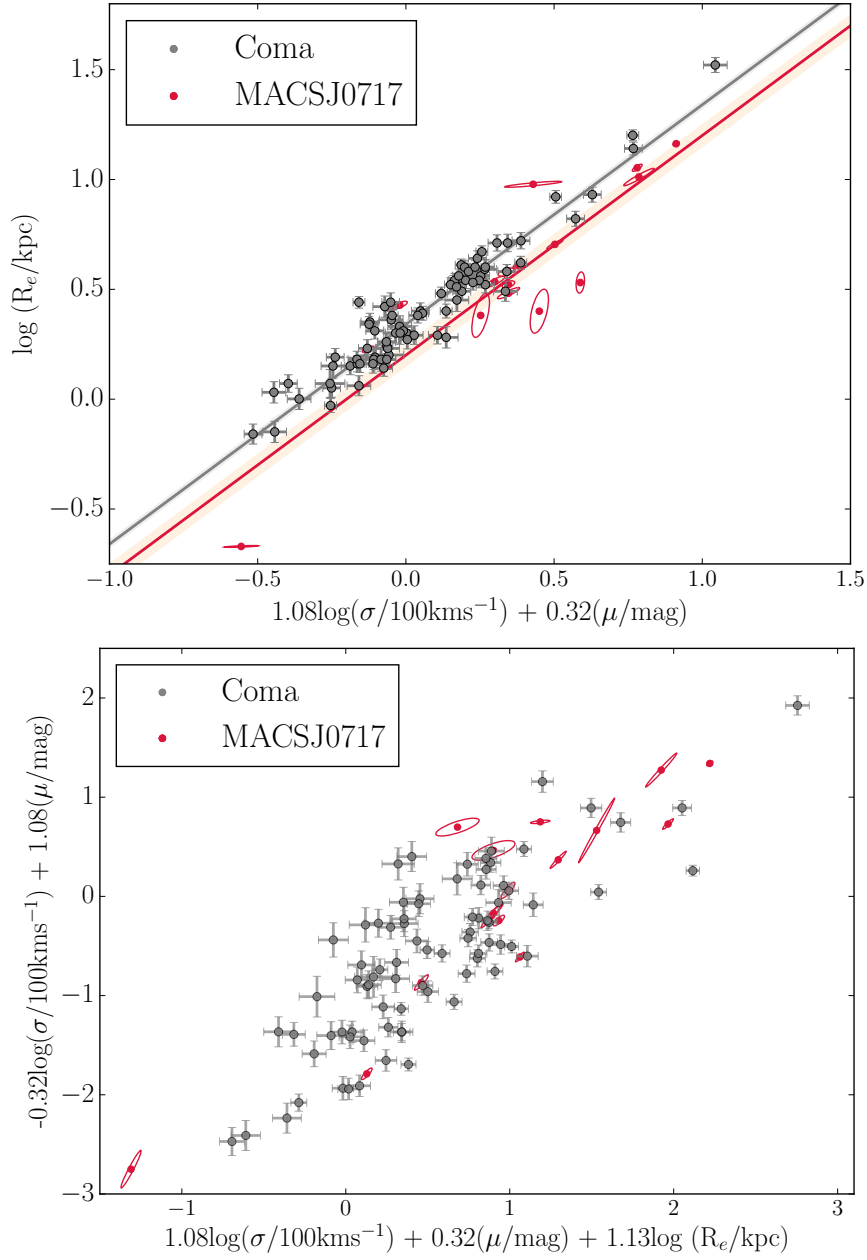


Figure 4.3: The fundamental planes of the Coma and MACSJ0717 galaxies, with σ in units of 100 km s^{-1} , $\langle \mu_e \rangle$ in magnitudes and R_e in kpc. The fundamental plane of MACSJ0717 is slightly offset from that of Coma, which we show to be the result of two effects: the evolution of the stellar populations and the evolution of the structure of the galaxies. Note that the surface brightnesses here are measured in the observed-frame F625W and r' filters for MACSJ0717 and Coma respectively, and have been corrected for cosmological dimming as $10 \log(1+z)$. Two complementary projections are shown in order to highlight the different regions occupied by the two galaxy populations. Top: An edge-on projection. Bottom: A face-on projection.

This is appropriate as it deals with the fact that the galaxies in each sample are almost certainly drawn from different intrinsic distributions, with the higher-redshift MACSJ0717 galaxies likely to be drawn from the most massive and luminous end of their population. More details on the model are provided in Kelly (2007), which presents the formalism, but essentially, we construct the likelihood for the data given a particular set of plane and Gaussian parameters, and obtain the posterior distribution using an MCMC exploration of the parameter space. The median values of the marginalised distributions are presented in Table 4.3, together with 16th and 84th percentile uncertainties, and the fitted fundamental plane is shown in both edge-on and face-on projections in Figure 4.3. We demonstrate that the MACSJ0717 galaxies are indeed drawn from a distribution with a higher mean velocity dispersion, with $v_{\log\sigma} = 0.31 \pm 0.03$ for MACSJ0717 and $v_{\log\sigma} = 0.21 \pm 0.02$ for Coma – consistent with the idea that the former extend to brighter magnitudes and higher masses. When we model the two planes simultaneously, we find the clusters to be offset, with $\gamma_z - \gamma_0 = -0.14 \pm 0.06$.

4.3.3 The $M_{dyn}/L - M_{dyn}$ relation

To construct the MLM relation, we calculate the dynamical mass-to-light ratio within R_e , M_{dyn}/L , in the $F625W$ band (MACSJ0717) and r band (Coma) using the virial estimator

$$(4.7) \quad \frac{M_{dyn}}{L}(< r_e) = \frac{\beta \sigma^2 R_e}{GL}$$

for 3D half-light radius r_e , adopting the best-fitting value of $\beta = 5$ presented in Cappellari et al. (2006), which was calibrated by comparing virial and Schwarzschild M_{dyn}/L estimates for a sample of 25 E and S0 galaxies (and also used a similar approach to measuring the effective radii and luminosities). We then construct a ‘dynamical’ mass

$$(4.8) \quad M_{dyn} = \frac{\beta \sigma^2 R_e}{G}$$

though we note that this is not intended to represent the total dynamical mass within any physically meaningful aperture (Equation 4.7 implies that $M_{dyn}(R < R_e)$ requires $\beta = 2.5$, whereas Wolf et al., 2010 advocate $\beta = 4$). We continue to use $\beta = 5$ in order to facilitate comparisons with other studies, and note that, given that we are interested in the offset between the two galaxy samples rather than the absolute relations, the choice of any constant β does not affect our conclusions (though this is not true if β varies across the plane).

We calculate the bandpass luminosities relative to that of the Sun (based on a redshifted CALSPEC solar spectrum) from the surface brightnesses using the equation

$$(4.9) \quad \mathcal{M}_{gal} - \mathcal{M}_{\odot} = -2.5 \log \left(\frac{L_{gal}}{L_{\odot}} \right)$$

where \mathcal{M} represents an absolute magnitude and $\mathcal{M}_{\odot}(z=0, r) = 4.58$ and $\mathcal{M}_{\odot}(z=0.545, F625W) = 5.17$, and construct the MLM relations for the Coma and MACSJ0717 galaxies, as can be seen in Figure 4.4. Again modelling the masses and mass-to-light ratios as being drawn from a multivariate Gaussian distribution, we infer both these underlying distributions and the slope, intercept and scatter of the linear relation

$$(4.10) \quad \log \frac{M_{dyn}}{L} = \alpha_{ML} \log M_{dyn} + \beta_{ML}.$$

where mass M_{dyn} and luminosity L are measured in units of $10^{10} M_{\odot}$ and L_{\odot} ; our inference is presented in Table 4.2. This time, modelling the clusters independently leads us to find marginally different slopes (in addition to an offset), with $\beta_{ML} = 0.25 \pm 0.02$ for Coma and $\beta_{ML} = 0.12 \pm 0.11$ for MACSJ0717, though the uncertainties of the MACSJ0717 relation are large, making them consistent at the 1σ level. The point here is that our high-redshift sample lacks the dynamic range that would be needed to robustly infer both the slope and intercept of the MLM relation; these two parameters suffer degeneracies, making them hard to constrain. When we then model the two populations together, we find that the Coma data dominate the fit to the slope – which is not surprising, given that we have ~ 6 times more galaxies in the latter – such that $\beta_{ML} = 0.25 \pm 0.02$; the MLM relation for MACSJ0717 then lies virtually on top of that of Coma, with $\beta_z - \beta_0 = -0.01 \pm 0.06$. We also confirm that the MACSJ0717 galaxies are drawn from a more massive distribution, with $\nu_{ML} = 1.24 \pm 0.15$ compared to $\nu_{ML} = 0.88 \pm 0.05$ for Coma. This is a selection effect that we would expect, as the former is at a higher redshift and can only observe the most massive end of the mass distribution. Nevertheless, Figure 4.5 suggests there are genuinely more high-mass galaxies in MACSJ0717 than in Coma – this is an interesting result that may be connected with the higher cluster mass of MACSJ0717 (see Section 4.5.4). Figure 4.4 also shows more generally that the two populations have similar trends between their dynamical masses and mass-to-light ratios, with the main difference being that the MACSJ0717 galaxies have higher masses. The results of this modelling are summarised in Table 4.2.

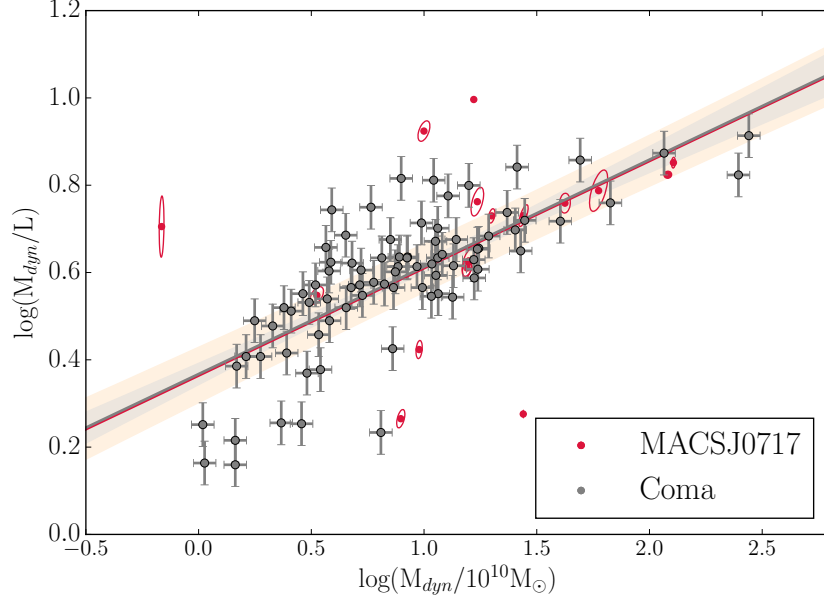


Figure 4.4: The MLM relation for the galaxies in Coma and MACSJ0717, with M_{dyn} in units of $10^{10} M_{\odot}$ and M_{dyn}/L in units of M_{\odot}/L_{\odot} . When the MACSJ0717 galaxies are constrained to be parallel to those of Coma, the relations lie virtually on top of each other (though note that the luminosity is measured in a different filter and at a different redshift for each cluster).

4.3.4 The $M_{\star} - \sigma$ and $M_{\star} - R_e$ relations

As discussed in the Introduction, any structural evolution of the galaxy population with redshift would also have an impact on the offset of the fundamental plane. We therefore attempt to measure the evolution in size and velocity dispersion between the two ETG samples using two independent methods. First, we fit the $M_{\star} - R_e$ and $M_{\star} - \sigma$ relations to infer the difference in size and velocity dispersion of galaxies of any given mass; that is the subject of this section. Later, we model the fundamental plane and MLM, $M_{\star} - R_e$ and $M_{\star} - \sigma$ relations simultaneously in order to infer the evolution in size, velocity dispersion and luminosity all at once. That is the topic of Section 4.4.

We use stellar masses here as opposed to the dynamical masses calculated in the previous section to avoid the obvious degeneracies between M_{dyn} , σ and R_e . We calculate stellar masses for Coma cluster and MACS0717 galaxies by comparing the $u - g$ and F475W-F625W colours, respectively, to the same colours calculated from the SSP models of Bruzual & Charlot (2003). In order to break the degeneracy between age and metallicity, we

make the assumption that galaxies on the red sequence are coeval to first-order and that the slope of the red sequence is driven by a systematic change in metallicity with luminosity (as found by Kodama & Arimoto, 1997). By assuming an *average* age of the red sequence, we interpolate between SSP models of fixed age and varying metallicity to convert the colour-magnitude relation of the red sequence into a metallicity-luminosity relation; this provides a metallicity for each galaxy, based on its luminosity and not its colour. With this metallicity, we use the observed colour to infer the age and stellar mass-to-light ratio M_\star/L of each galaxy. In practice, for each galaxy we interpolate between the two SSPs that bracket the metallicity derived from the CMR to calculate the age and M_\star/L . We do not apply any luminosity weighting corrections as these are deemed second order. Nor do we limit the derived ages to be younger than the age of the Universe (although this is not a significant issue).

For young star forming galaxies, M_\star/L is far smaller, and the luminosity far greater, than that of an old passive galaxy of the same mass, leading us to dramatically overestimate of the metallicity from the luminosity-metallicity relation we derived above for red sequence galaxies. However, at young ages ($< 1\text{Gyr}$), colour is primarily determined by age, not metallicity. Thus curves of colour versus M_\star/L for different metallicities only diverge at old ages; at younger ages, the curves converge and the dominant factor in determining M_\star/L is age; metallicity has virtually no effect. Thus for luminous blue star forming galaxies, although our luminosity-metallicity relation overestimates the metallicity, the M_\star/L and stellar masses remain accurate. Hence we find no need to iterate the estimation of the metallicity once the age and M_\star/L (the real parameters of interest) have been calculated.

Adopting this method, we calculate the M_\star/L for the Coma galaxies assuming red sequence ages of 8, 10, & 12 Gyrs. We then adopt the average of the M_\star/L values and use the scatter from the variation of input age as the formal error on the M_\star/L values. We find Coma cluster galaxies to be around solar metallicity and (by construction) around 10 Gyrs old. When calculating the M_\star/L of the galaxies in MACS0717, we adopt red sequence ages of (8, 10, 12) - 5.3 Gyrs. We find that the MACSJ0717 galaxies are slightly higher metallicity (around $1.5 \times$ solar) than Coma galaxies (around solar). Note that although the derived stellar masses depend slightly on the initial age assumed for Coma (and for MACS0717, forced to be 5.3 Gyrs younger), the relative ratio between the Coma and MACS0717 masses is almost constant if we assume an age > 8 Gyrs for Coma. For ages below 8 Gyrs, the stellar masses for MACSJ0717 galaxies drop rapidly compared to the stellar masses for Coma galaxies (to the extent that if we assume an age of 6 Gyrs for Coma, the average stellar masses in both clusters become the same). But such young ages (implying ages of $< 1\text{Gyr}$ for the MACSJ0717 stellar populations) are unrealistic and ruled

out by both the CMR and fundamental plane or MLM results.

The stellar masses are included in Table 4.1; we note that the stellar masses reveal the general mass differences of the two samples even more clearly than the dynamical masses of the previous section, with almost all the high- z galaxies containing more stellar mass than nearly all the low- z galaxies. As discussed previously, this indicates a genuine excess of high-mass galaxies in MACSJ0717 relative to Coma.

We assume the size and velocity dispersion to follow power laws in the total mass and fit

$$(4.11) \quad \log R_e = \alpha_{MR} \log M_\star + \beta_{MR}$$

and

$$(4.12) \quad \log \sigma = \alpha_{MS} \log M_\star + \beta_{MS}$$

using the same formalism as in earlier sections in which the slope, intercept and intrinsic scatter and inferred together with the properties of the underlying Gaussian distribution of $\log M_\star$. The limited dynamic range in mass of each population makes it difficult to make meaningful inference on both the slope and intercept of these relations (especially for the size-mass relation) and break the strong degeneracy that exists between them: we therefore fix the slope of the size-mass relation to $\alpha_{MR} = 0.56$, as found in Shen et al. (2003), and that of the sigma-mass relation to $\alpha_{MS} = 0.23$ as in Saglia et al. (2010), both of which were measured using significantly bigger galaxy samples. Our inference is shown in Figure 4.5 and demonstrates that the two populations look extremely similar in both respects. This implies that a very small amount of evolution has taken place between the $z \sim 0.5$ and $z \sim 0$ galaxies.

We also compare our $M_\star - R_e$ and $M_\star - \sigma$ relations with those of Saglia et al. (2010) – in which these relations were constructed for galaxies from 26 clusters out to redshifts $z \sim 0.9$ – and find that we are consistent with both at the 2σ level. With regard to the slightly poorer agreement between the $M_\star - \sigma$ relations relative to the $M_\star - R_e$ relations, we note that MACSJ0717 has a high velocity dispersion $\sigma_{cluster} = 1660^{+120}_{-130} \text{ kms}^{-1}$ (Ebeling et al., 2007) relative to the mean velocity dispersion $\langle \sigma_{cluster} \rangle = 525 \pm 210 \text{ kms}^{-1}$ of the EDisCS clusters used in that study. This may be evidence for the more rapid evolution of galaxies in denser environments, such that the galaxy velocity dispersions in higher-mass clusters at $z = 0.5$ more closely resemble those of $z = 0$ galaxies than do those in lower-mass clusters at $z = 0.5$. We discuss this further in Section 4.5.2, though we cannot make any strong claims on the basis of the data used in this study.

Following van der Wel et al. (2008) and Saglia et al. (2010), we relate the offsets between the clusters to measure the evolution as a function of redshift:

$$(4.13) \quad \beta_{MR,z} - \beta_{MR,0} = \xi \log(1+z)$$

and

$$(4.14) \quad \beta_{MS,z} - \beta_{MS,0} = \eta \log(1+z)$$

and find $\xi = -0.37 \pm 0.39$, $\eta = 0.06 \pm 0.28$ as summarised in Table 4.2. We note that the evolution we find here is weaker than the $\xi = -0.98 \pm 0.11$ found by van der Wel et al. (2008) for field ellipticals, but consistent with the $\xi = -0.53 \pm 0.04$ inferred by Delaye et al. (2014) for cluster ellipticals, and moreover that the evolution in both size and velocity dispersion is consistent with zero (see also Saglia et al., 2010; Jørgensen et al., 2014; Saracco et al., 2014, for indications of mild structural evolution of cluster ETGs). The implication is that only a small amount of evolution has taken place in these galaxies between $z = 0.545$ and the present day. We use these in Section 4.4 to account for the effects of size evolution in the fundamental plane and MLM relations.

4.4 Stellar Population Models

The key idea of using the changes in our scaling relations to understand the underlying stellar populations is that the scaling relations themselves are simply a convenient way of characterising the observable properties of galaxies - in our case, this means luminosities, colours, sizes, velocity dispersions - and that these properties are governed not only by a galaxy's mass structure, but also by the stars it contains. We can then use stellar population models, under particular, astrophysically-motivated assumptions regarding the age, metallicity, IMF and star formation history of the population, to connect the changes we observe to the evolution of the stellar populations. The assumptions we make as to the star formation histories in the CMR and the fundamental plane and M_{dyn}/L analyses are different and are simplifications of the real, much more complex and extended processes that we know ETGs are subject to: however, both are motivated by the data, and can be interpreted together to provide a fuller picture of these galaxies' evolution. The other main assumption we make here is that the ETGs in MACSJ0717 are directly comparable to those in Coma, such that the former represent the Coma population at an earlier stage in their evolution.

Historically, the differences due to the redshifts of the two galaxy populations and the filters in which they have been observed have been accounted for by *correcting the data*;

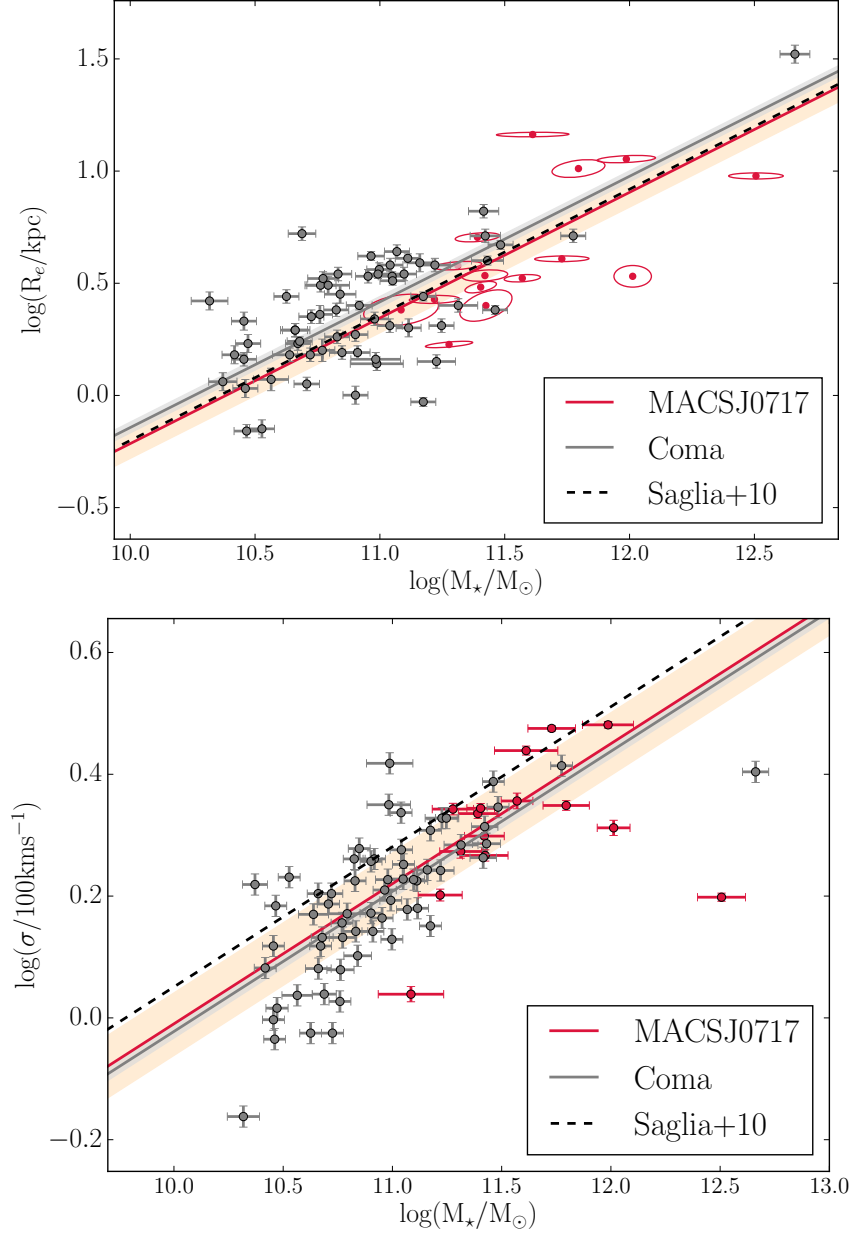


Figure 4.5: Mass-size and mass-velocity-dispersion relations for the two clusters, with stellar masses in units $10^{10}M_\odot$, sizes in kpc and velocity dispersions in units of 100 km s^{-1} . In both cases, the scaling relations of the two clusters are consistent with being parallel but offset, with the high-redshift galaxies being both marginally smaller and having marginally higher velocity dispersions. The dashed lines show the relations given in Saglia et al. (2010), evaluated at $z = 0.545$. Both are consistent with our MACSJ0717 sample within 2σ (though we do not show the uncertainties on the Saglia et al., 2010 relations here).

specifically, by applying K-corrections to the data and then comparing the fundamental plane zeropoints as if both galaxy populations had been observed at the same redshift and with the same filter. However, this requires some galaxy ‘template’ to be chosen and assumptions to be made regarding the spectral energy distribution (SED) and age of the galaxies, and therefore introduces significant uncertainty and possible bias. We therefore refrain from doing this, and, rather than correcting the data, *we entirely forward-model the observations*. In a development of the methods introduced in Houghton et al. (2012), we use the BC03 stellar population models, assuming a Salpeter IMF (based on evidence that massive ETGs may have IMFs that more heavier than the Milky-Way-like Chabrier IMF, e.g. Auger et al., 2010b) and solar metallicity, and entirely forward-model the data by evaluating the stellar population models in the same redshifted filters as those with which the latter were observed (though we do subtract the cosmological dimming term $10\log(1+z)$ from the surface brightnesses, as stated in Section 4.2.2). This removes the need to apply any ad-hoc colour or bandpass ‘corrections’ to the data (see e.g. Hogg et al., 2002) based on assumptions about the SED. Any further assumptions that we make regarding the stellar populations in the case of specific scaling relations are explained in the relevant section.

4.4.1 Luminosity evolution from the CMR

To constrain the stellar ages from the evolution in the CMR scatter, we assume a star formation history in which each galaxy comprises an SSP, but allow for a spread in SSP ages between galaxies. This allows us to write the colour scatter as a Taylor expansion

$$(4.15) \quad \frac{d\text{col}}{dt} \approx \frac{\sigma_{CMR}}{\sigma_{age}} \approx 3.5 \frac{\sigma_{CMR}}{\Delta t}$$

where $\frac{d\text{col}}{dt}$ is the rate of change of colour at the mean stellar age of the galaxies, σ_{CMR} is the intrinsic colour scatter measured from the CMR, σ_{age} is the intrinsic scatter in stellar age between galaxies (assuming a Gaussian distribution) which translates to the equivalent width of $\Delta t \approx 3.5\sigma_{age}$ of a population of galaxies that form their SSPs uniformly between t_{start} and t_{stop} . Relating t_{start} and t_{stop} by the ratio b of the SSP formation period to the total time available,

$$(4.16) \quad \Delta t = t_{stop} - t_{start} = b t_{stop},$$

we can rewrite the mean formation time $t_f = \frac{1}{2}(t_{start} + t_{stop})$ in terms of Δt and so derive an equation for the evolution in colour of a galaxy’s stars:

$$(4.17) \quad \frac{d\text{col}}{dt} = 3.5\sigma_{CMR} \frac{b^{-1} - \frac{1}{2}}{t_f}$$

(see also Bower et al., 1992, 1998; Houghton et al., 2012). Note that this model has all the stars in a single galaxy forming simultaneously, but distributes the formation times for different galaxies uniformly with a dispersion b and mean age t_f . Thus $b = 0$ corresponds to a cluster whose galaxies all formed at once, whereas $b = 1$ allows the cluster galaxies to have formed their stars from the beginning of the Universe until t_{stop} .

Given that we have inferred the intrinsic scatter σ_{CMR} for the CMRs of the two clusters, and that we can use stellar population models to calculate the rate of change of colour as a function of stellar age, we can thus infer the stellar age dispersion b and mean stellar age $13.6 - t_f$ Gyr of the cluster galaxies. We do this by constructing a chi-squared likelihood from Equation 4.15, accounting for the uncertainties in σ_{CMR} that are given in Table 4.2. As shown in Figure 4.6 (left), we find a mean age $9.44^{+0.46}_{-0.57}$ Gyr and dispersion $b = 0.83^{+0.11}_{-0.12}$, indicating that the stellar populations are fairly old but formed with a significant dispersion ~ 3 Gyr. Note that this is a strong constraint compared to the lower limits that have been previously obtained from the CMR (Bower et al., 1992, 1998; Houghton et al., 2012); this is mainly due to the fact that we have measured the intrinsic scatter for two clusters rather than one, and can therefore break the degeneracy between b and t_f , which would otherwise be unconstrained. The strength of our constraint relative to Houghton et al. (2012), which also used two clusters, is due to the wider redshift separation of our clusters. The constraint obtained here could therefore be further improved by the addition of further, higher-redshift clusters.

4.4.2 Luminosity evolution from the fundamental plane

We have observed that there is a change in the zero-point of the fundamental plane with time. This could be due to an evolution in any or all of the plane variables - indeed, at fixed mass, all are understood to change with redshift. Similarly to Saglia et al. (2010) and van der Wel et al. (2008), we derive a general expression relating the fundamental plane offset to the evolution of the fundamental plane variables, and apply this, first under the assumption of no size evolution, and then using a parameterisation for size evolution which we constrain directly from our data. In contrast to the CMR analysis, we now assume that all the stars in all the galaxies formed in a single event, with one SSP characterising the whole cluster. We can then compare the luminosity evolution we observe with the predictions of stellar population models to obtain a further estimate of the stellar age.

By requiring that the fundamental plane of MACSJ0717 be parallel to that of Coma, we have been able to measure its displacement, $\Delta\gamma_{FP} = \gamma_{FP,z} - \gamma_{FP,0} = -0.14 \pm 0.06$, in the direction of $\log R_e$, where subscripts x_0 and x_z denote quantities measured with respect to

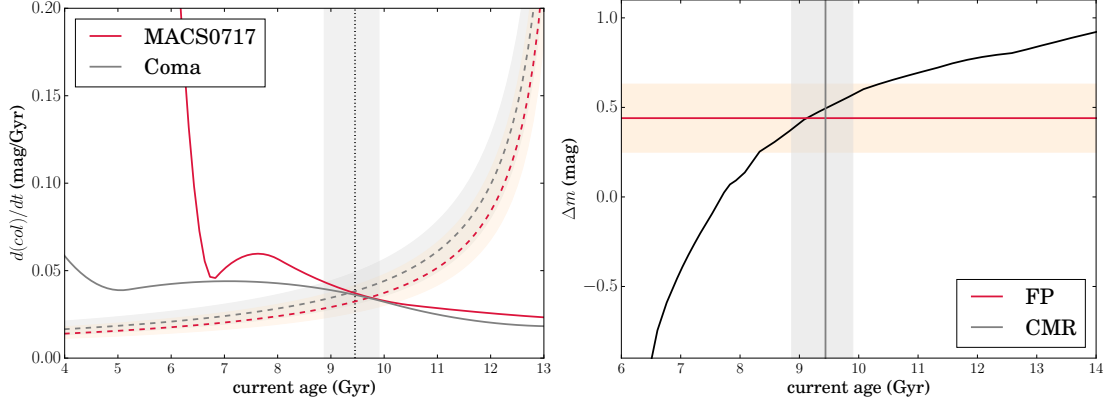


Figure 4.6: Left: Inferring the mean stellar age and age dispersion of the cluster galaxies from the evolution of the CMR’s intrinsic scatter. Assuming each galaxy to consist of an SSP, with a spread formation times between galaxies, we compare the implied rate of change of colour with that predicted by SSP models. The solid curves show the rate of change of colour that for a solar metallicity SSP with a Salpeter IMF in the BC03 models as a function of the current age of the stars; the dashed curves show the right-hand side of Equation 4.16 for the value $b = 0.83^{+0.11}_{-0.12}$ that we infer, and the vertical black line shows the inferred mean age $9.44^{+0.46}_{-0.57}$ Gyr. Right: SSP prediction for the evolution of the observed magnitude difference between the Coma and MACSJ0717 galaxies as a function of cluster age. Overplotted in red is the offset inferred from the evolution of the fundamental plane, as described in Section 4.4.2; the grey vertical line shows the inference on age from the CMR, which agrees remarkably well. In both cases, the shaded regions show the 1σ uncertainties.

the low- and high-redshift clusters respectively, as before. Given the fundamental plane equation and the construction of the surface brightness as $\langle\mu_e\rangle = -2.5 \log(\frac{L}{2\pi R_e^2})$, we have

$$(4.18) \quad \gamma_z - \gamma_0 = \Delta \log R_e - \alpha \Delta \log \sigma - \beta \Delta \langle\mu_e\rangle$$

where we have dropped the subscript x_{FP} for clarity, and $\Delta X = X_z - X_0$, i.e. $\Delta X > 0$ means that X has decreased between redshift z and today. This translates into a magnitude evolution Δm

$$(4.19) \quad \Delta m = \frac{(1 - 5\beta)\Delta \log R_e - \alpha \Delta \log \sigma - \Delta \gamma}{\beta}.$$

In Section 4.4.5, we use the full machinery of Equation 4.19 to both investigate and treat the effects of size evolution. Here, though, we proceed under the first approximation that R_e and σ are constant with redshift, in which case Equation 4.19 simplifies dramatically.

The measured change between γ_z and γ_0 translates to an evolution in magnitude of $\Delta m = 0.44 \pm 0.10$ mag (in the different filters), which we include in Table 4.4. This implies

that, for a particular position on the fundamental plane, the stars in a MACSJ0717 galaxy are actually *dimmer* than those in Coma. While, at first glance, this appears contrary to our expectation of ageing stellar populations, we emphasise that the high- and low-redshift filters are *not* matched, meaning that each is being sampled in a different region of the spectrum. The purpose of our stellar population model comparison is to account for these filter effects in addition to the effects of the intrinsic luminosity evolution of the population.

We use the BC03 stellar population models to interpret this, again assuming an SSP with solar metallicity and a Salpeter IMF. As explained at the beginning of Section 4.4, we take account of both the different filters used and the age difference by modelling the colour $\Delta m = F625W(z = 0.545, T - 5.33) - r(z = 0, T)$, where z is the redshift and T is the age of the Universe in Gyr. The evolution of Δm in this setup is shown in Figure 4.6 (right), with the magnitude offset $\Delta m = 0.44 \pm 0.19$ determined from the fundamental plane overplotted in red, along with its 1σ upper and lower bounds. The intersection of this measured offset with the model lines gives an estimate for the age as $9.12^{+1.22}_{-0.80}$ Gyr, and is in good agreement with the constraints from the CMR that were obtained in the previous section, plotted in grey on the figure. This is old, and implies that the stellar populations are already highly evolved by the time we observed them in MACSJ0717.

4.4.3 Luminosity evolution from the $M_{dyn}/L - M_{dyn}$ relation

As for the fundamental plane, we assume that both populations can be described by the same MLM slope (the individual cluster slopes are consistent at the 1σ level) and, under the same assumption of one universal SSP, we use our simultaneous fit to the two clusters and compare the offset between them with that predicted by SSP models in order to make another estimate of the stellar age. Here, given that the mass scales as $M \sim \sigma^2 R_e$, we have

$$(4.20) \quad \Delta \log(M/L) = 2\Delta \log \sigma + \Delta \log R_e - \Delta \log L$$

which can be related to the offset using Equation 4.9 to give

$$(4.21) \quad \Delta m = \Delta \mathcal{M}_\odot + 2.5[\Delta \beta_{ML} + (\alpha_{ML} - 1)(2\Delta \log \sigma + \Delta \log R_e)]$$

where $\Delta \mathcal{M}_\odot$ is the difference in absolute solar magnitude in the (blueshifted) filters.

As with the fundamental plane, we initially assume zero structural evolution, and calculate the magnitude offset and corresponding stellar age. From our best fit to the two cluster populations, we then find $\Delta m = 0.57 \pm 0.18$, consistent with our earlier fundamental plane result and included in Table 4.4. This implies a stellar age of $9.89^{+1.73}_{-0.98}$ Gyr.

4.4.4 Combining size and luminosity evolution

To apply our inference on size evolution from the mass-size and mass-velocity dispersion relations of Section 4.3.4 to our scaling relations using Equations 16 and 18, we define

$$(4.22) \quad \Delta \log R_e = \xi \log(1+z)$$

and

$$(4.23) \quad \Delta \log \sigma = \eta \log(1+z),$$

using the values of ξ and η as defined in Equations 4.13 and 4.14 and tabulated in Tables 4.2 and 4.3.

Using our calculated values for ξ and η alongside equations 16 and 18, we can now calculate the magnitude evolution Δm that must have taken place between the two clusters according to the fundamental plane and the MLM relation. We can write the magnitude offset as

$$(4.24) \quad \Delta m = \frac{\log(1+z) \left[(1-5\beta)\xi - \alpha\eta \right] - \Delta\gamma}{\beta}$$

for the fundamental plane, and

$$(4.25) \quad \Delta m = \Delta \mathcal{M}_\odot + 2.5 \left[\Delta\beta + (\alpha-1)(2\eta + \xi) \log(1+z) \right]$$

for the MLM relation.

In Table 4.4, we summarise η , ξ , Δm and the corresponding stellar age, according to our stellar population models, both with and without allowing for structural evolution. The final two columns of that table give the implied magnitude evolution in the rest-frame Johnson B and V bands, assuming the stellar age that has been inferred in each case. We note that we do not attempt to account for progenitor bias that may arise from the fact that relatively young galaxies included in the local sample would be missing at higher redshifts, where they would not yet appear passive; we do not have the data to constrain it here. However, Valentinuzzi et al. (2010b) and Saglia et al. (2010) investigated this for their larger samples of field galaxies at comparable redshifts, and found the effect to be small. We therefore assume the same to be true for our cluster galaxy samples. The MLM relation generally implies ages that are between 0.5 and 1 Gyr older than the fundamental plane, though the results are formally consistent and imply a stellar age of ~ 10 Gyr.

parameter	CMR	FP	FP (parallel)	MLM	MLM (parallel)	$M_\star - \sigma$ (parallel)	$M_\star - R_e$ (parallel)	ALL
α_{CMR}	-0.03 ± 0.01	—	—	—	—	—	—	—
β_{CMR}	1.17 ± 0.01	—	—	—	—	—	—	—
σ_{CMR}	0.055 ± 0.009	—	—	—	—	—	—	—
α_{FP}	—	1.15 ± 0.43	1.08 ± 0.07	—	—	—	—	1.10 ± 0.18
β_{FP}	—	0.34 ± 0.05	0.32 ± 0.01	—	—	—	—	0.34 ± 0.03
γ_{FP}	—	0.18 ± 0.15	0.20 ± 0.05	—	—	—	—	—
$\nu_{FP, \log \sigma}$	—	0.31 ± 0.03	0.31 ± 0.03	—	—	—	—	0.31 ± 0.03
ν_{FP, μ_e}	—	0.10 ± 0.28	0.08 ± 0.30	—	—	—	—	0.10 ± 0.29
$\tau_{FP, \log \sigma}$	—	0.12 ± 0.03	0.12 ± 0.03	—	—	—	—	0.12 ± 0.03
τ_{FP, μ_e}	—	1.13 ± 0.25	1.15 ± 0.25	—	—	—	—	1.15 ± 0.26
ρ	—	0.11 ± 0.25	0.12 ± 0.27	—	—	—	—	0.12 ± 0.26
σ_{FP}	—	0.17 ± 0.05	0.16 ± 0.04	—	—	—	—	0.12 ± 0.06
α_{ML}	—	—	—	0.12 ± 0.11	0.25 ± 0.02	—	—	0.25 ± 0.04
β_{ML}	—	—	—	0.52 ± 0.15	0.36 ± 0.06	—	—	—
ν_{ML}	—	—	—	1.24 ± 0.15	1.24 ± 0.15	—	—	1.25 ± 0.16
τ_{ML}	—	—	—	0.60 ± 0.14	0.59 ± 0.14	—	—	0.60 ± 0.14
σ_{ML}	—	—	—	0.22 ± 0.05	0.22 ± 0.05	—	—	0.22 ± 0.05
β_{MR}	—	—	—	—	—	—	-0.21 ± 0.07	—
ν_{MR}	—	—	—	—	—	—	1.39 ± 0.22	1.39 ± 0.15
τ_{MR}	—	—	—	—	—	—	0.90 ± 0.20	0.86 ± 0.13
σ_{MR}	—	—	—	—	—	—	0.24 ± 0.06	0.24 ± 0.06
β_{MS}	—	—	—	—	—	-0.01 ± 0.05	—	—

ν_{MS}	—	—	—	—	—	1.39 ± 0.23	—	1.39 ± 0.15
τ_{MS}	—	—	—	—	—	0.92 ± 0.20	—	0.86 ± 0.13
σ_{MS}	—	—	—	—	—	0.20 ± 0.05	—	0.20 ± 0.04
ξ	—	—	—	—	—	—	-0.37 ± 0.39	-0.40 ± 0.32
η	—	—	—	—	—	0.06 ± 0.28	—	0.09 ± 0.27
Δm	—	—	—	—	—	—	—	0.59 ± 0.26

Table 4.2: The inferred parameters for the CMR, fundamental plane, MLM, MR and MS relations of MACSJ0717, modelled as described in Section 4.3.

parameter	CMR	FP	FP (parallel)	MLM	MLM (parallel)	$M_{\star} - \sigma$ (parallel)	$M_{\star} - R_e$ (parallel)	ALL
α_{CMR}	-0.11 ± 0.01	—	—	—	—	—	—	—
β_{CMR}	-0.82 ± 0.01	—	—	—	—	—	—	—
σ_{CMR}	0.065 ± 0.005	—	—	—	—	—	—	—
α_{FP}	—	1.09 ± 0.07	1.08 ± 0.07	—	—	—	—	1.10 ± 0.18
β_{FP}	—	0.32 ± 0.01	0.32 ± 0.01	—	—	—	—	0.34 ± 0.03
γ_{FP}	—	0.34 ± 0.02	0.34 ± 0.02	—	—	—	—	0.35 ± 0.04
$\nu_{FP, \log \sigma}$	—	0.21 ± 0.02	0.21 ± 0.02	—	—	—	—	0.19 ± 0.02
ν_{FP, μ_e}	—	-0.48 ± 0.09	-0.48 ± 0.09	—	—	—	—	-0.53 ± 0.10
$\tau_{FP, \log \sigma}$	—	0.14 ± 0.01	0.14 ± 0.01	—	—	—	—	0.12 ± 0.01
τ_{FP, μ_e}	—	0.78 ± 0.07	0.78 ± 0.06	—	—	—	—	0.82 ± 0.08
ρ	—	-0.09 ± 0.11	-0.09 ± 0.11	—	—	—	—	-0.24 ± 0.13
σ_{FP}	—	0.07 ± 0.01	0.08 ± 0.01	—	—	—	—	0.05 ± 0.03
α_{ML}	—	—	—	0.25 ± 0.04	0.25 ± 0.02	—	—	0.25 ± 0.04

β_{ML}	—	—	—	0.36 ± 0.04	0.37 ± 0.02	—	—	0.37 ± 0.03
ν_{ML}	—	—	—	0.88 ± 0.05	0.88 ± 0.06	—	—	0.82 ± 0.05
τ_{ML}	—	—	—	0.49 ± 0.04	0.48 ± 0.04	—	—	0.39 ± 0.04
σ_{ML}	—	—	—	0.07 ± 0.01	0.10 ± 0.01	—	—	0.10 ± 0.01
α_{MR}	—	—	—	—	—	—	—	—
β_{MR}	—	—	—	—	—	—	-0.14 ± 0.03	-0.14 ± 0.02
ν_{MR}	—	—	—	—	—	—	0.93 ± 0.05	0.93 ± 0.04
τ_{MR}	—	—	—	—	—	—	0.38 ± 0.04	0.38 ± 0.03
σ_{MR}	—	—	—	—	—	—	0.19 ± 0.02	0.19 ± 0.02
α_{MS}	—	—	—	—	—	—	—	—
β_{MS}	—	—	—	—	—	-0.02 ± 0.01	—	-0.02 ± 0.01
ν_{MS}	—	—	—	—	—	0.93 ± 0.05	—	0.93 ± 0.04
τ_{MS}	—	—	—	—	—	0.38 ± 0.04	—	0.38 ± 0.03
σ_{MS}	—	—	—	—	—	0.08 ± 0.01	—	0.08 ± 0.01
ξ	—	—	—	—	—	—	-0.37 ± 0.39	-0.40 ± 0.32
η	—	—	—	—	—	0.06 ± 0.28	—	0.09 ± 0.27
Δm	—	—	—	—	—	—	—	0.59 ± 0.26

Table 4.3: The inferred parameters for the CMR, fundamental plane, MLM, MR and MS relations of Coma, modelled as described in Section 4.3.

scaling relation	size evolution corrected?	ξ	η	observed-frame magnitude evolution / mag	age / Gyr	rest-frame <i>B</i> -band magnitude evolution / mag	rest-frame <i>V</i> -band magnitude evolution / mag
CMR	–	–	–	–	$9.44^{+0.46}_{-0.57}$	–	–
FP	N	–	–	0.44 ± 0.19	$9.12^{+1.22}_{-0.80}$	$-0.86^{+0.15}_{-0.16}$	$-0.78^{+0.14}_{-0.16}$
FP	Y	-0.37 ± 0.39	0.06 ± 0.28	0.54 ± 0.29	$9.72^{+3.17}_{-1.40}$	$-0.78^{+0.21}_{-0.24}$	$-0.69^{+0.18}_{-0.23}$
MLM	N	–	–	0.57 ± 0.18	$9.89^{+1.73}_{-0.98}$	$-0.75^{+0.14}_{-0.15}$	$-0.67^{+0.12}_{-0.15}$
MLM	Y	-0.37 ± 0.39	0.06 ± 0.28	0.66 ± 0.19	$10.64^{+2.50}_{-1.34}$	$-0.69^{+0.13}_{-0.15}$	$-0.61^{+0.11}_{-0.14}$
MLM & FP	Y	-0.40 ± 0.32	0.09 ± 0.27	0.59 ± 0.26	$10.00^{+3.14}_{-1.32}$	$-0.74^{+0.19}_{-0.09}$	$-0.65^{+0.15}_{-0.08}$

Table 4.4: Inferences on the size, magnitude and velocity dispersion evolution from the fundamental plane, MLM and $M - R_e$ and $M - \sigma$ relations and the CMR, and the implied formation times of the stellar populations. All models recover a stellar age ~ 10 Gyr, though the MLM relation implies slightly larger ages than the fundamental plane and the joint analysis. The final two columns provide the implied magnitude evolution in the rest-frame Johnson B and V bands, according to the ages given in the sixth column.

4.4.5 Combining scaling relations: inferring size and luminosity evolution

In the previous sections, we analysed each scaling relation separately, using the size evolution measured from the MR (i.e. $M_\star - R_e$) and MS (i.e. $M_\star - \sigma$) relations to infer the evolution in luminosity and so the stellar age. Here, we model the fundamental plane and the MLM, MR and MS relations simultaneously, using both clusters so as to infer not only the scaling relation parameters and the underlying (Gaussian) distributions as before, but also to infer the magnitude, size and dynamical evolution between $z = 0.545$ and the present day. This has a number of significant advantages, including that (a) it ensures that the inferred scaling relations are all consistent, (b) it allows us to infer the physical parameters ξ , η and Δm (which were previously calculated after the modelling) in addition to those describing the scaling relations themselves and (c) it fully explores degeneracies between the physical parameters, as is not possible to do when they are calculated post-modelling.

To do this, we use the same formalism as before. We assume the MLM relations for the two clusters can be described according to the following equations

$$(4.26) \quad \log \frac{M_{dyn}}{L} = \alpha_{ML} \log M_{dyn} + \beta_{ML}(z)$$

where the intercept for the low-redshift cluster is

$$(4.27) \quad \beta_{ML}(0) = \beta_{ML}$$

and for MACSJ0717 we now explicitly account for size, magnitude and velocity dispersion evolution with

$$(4.28) \quad \beta_{ML}(z) = \beta_{ML} + 0.4(\Delta m - \Delta \mathcal{M}_\odot) + (1 - \alpha_{ML})(2\eta + \xi) \log(1 + z).$$

As before, we take the slope α_{ML} to be the same for both clusters. Also as before, each cluster has a distribution in $\log M_{dyn}$ given by a normal distribution with mean ν_{ML} and variance τ_{ML}^2 .

We also assume the fundamental planes for the two clusters can be described according to

$$(4.29) \quad \log R_e = \alpha_{FP} \log \sigma + \beta_{FP} \langle \mu_e \rangle + \gamma_{FP}(z)$$

where the intercept of the low-redshift cluster is

$$(4.30) \quad \gamma_{FP}(0) = \gamma_{FP}$$

and for MACSJ0717

$$(4.31) \quad \gamma_{FP}(z) = \gamma_{FP} + \left[\xi(1 - 5\beta_{FP}) - \alpha\eta \right] \log(1+z) - \beta_{FP}\Delta m.$$

Again, the slopes α_{FP} and β_{FP} are the same for both clusters, and the independent variables $\log\sigma$ and $\langle\mu_e\rangle$ are drawn from multivariate normal distributions as described in Section 4.3.2.

Finally, we include the MR and MS relations, retaining our definition of η and ξ as being measured at constant stellar mass, to give

$$(4.32) \quad \beta_{MR}(z) = \beta_{MR,0} + \xi \log(1+z)$$

and

$$(4.33) \quad \beta_{MS}(z) = \beta_{MS,0} + \eta \log(1+z)$$

for the MR and MS relations respectively, with $\log M_\star$ being drawn from a normal distribution with mean ν_{M_\star} and variance $\tau_{M_\star}^2$. This model now has the advantage of allowing us to infer the amount of size evolution and magnitude evolution that best describe our whole dataset, and guarantees that all four scaling relations are treated in a consistent way. It also sidesteps some of the potential dangers of our earlier method for constraining the structural evolution, as it does not assume that the stellar mass remains constant with redshift.

Our results are summarised in Table 4.2: encouragingly, the parameters of the fundamental plane and MLM relations are consistent with those inferred in our previous, simpler models. However, we now have additional constraints on the evolution of R_e , σ and the luminosity. Although ξ , η and Δm have degeneracies within each relation, the modelling of all four relations at once breaks this degeneracy and we are able to infer $\Delta m = 0.59 \pm 0.26$, $\xi = -0.40 \pm 0.32$ and $\eta = 0.09 \pm 0.27$. The fact that the uncertainties on these parameters are comparable to – and, in a number of cases, smaller than – the uncertainties on the same parameters when each scaling relation is modelled separately, indicates that the degeneracies are not significant, and that the scheme we have set up is indeed internally consistent. These results correspond to a stellar age $10.00^{+3.14}_{-1.32}$ Gyr.

4.5 Discussion

4.5.1 Old, passively evolving stellar populations

The fundamental plane and its evolution with redshift contain a wealth of information about ETG formation and evolution; however, in order to extract this information meaningfully, it is important to understand the contributions due to different processes – that is, the luminosity evolution of the stellar populations and the structural evolution of the galaxies themselves – and to find a way to disentangle them. In this work, we have combined the fundamental plane with a number of other scaling relations in order to break these degeneracies and make inference on both the luminous and structural evolution. We are now in a position to tie together what we have found.

Initially, we used the evolution of the intrinsic scatter of the CMR to infer the mean stellar age, allowing for some dispersion. There, we assumed a star formation history in which each galaxy is composed of an SSP with some dispersion in age across the galaxy population, and used the small evolution in the intrinsic scatter to infer a mean age $9.44^{+0.46}_{-0.57}$ Gyr and a dispersion of ~ 3 Gyr. This dispersion is significant but still implies some coordination in the star formation times of the different galaxies; together with the small intrinsic scatter of the fundamental plane and the MLM relation, this justified our treatment of the latter assuming that all the galaxies' stars formed in a single burst. We then modelled the fundamental plane, MLM, MS and MR relations in two ways: first, treating each separately, constraining the slopes to be parallel for the two clusters and using the offsets between them to measure the evolution in size, velocity dispersion and luminosity and hence the stellar age; second, by requiring all four relations to have evolved in a consistent way with regard to the structures and luminosities of the galaxies. In both cases, we find very clearly that only a small amount of evolution has taken place, with the high-redshift galaxies only marginally smaller and with marginally higher velocity dispersions than the Coma galaxies, and the luminosity evolution consistent with the passive fading of old populations.

In Table 4.4, we present the magnitude evolution that was inferred in each case, both in the observed-frame r' (Coma) and F625W (MACSJ0717) filters and the rest-frame U and V -band filters (though we note that the latter are more uncertain due to assumptions made in calculating K-corrections). In the joint analysis, we find $\Delta m_B = -0.74^{+0.19}_{-0.09}$ mag, or equivalently, $\Delta \log M/L_B = -0.30^{+0.08}_{-0.04} = -0.55^{+0.15}_{-0.07} z = -1.59^{+0.42}_{-0.21} \log(1+z)$. This is consistent with the findings of Saglia et al. (2010) and Holden et al. (2010), indicating that the stellar populations in these ETGs have been evolving passively. Our inferred mean age of $10.00^{+3.14}_{-1.32}$ Gyr, corresponding to a formation redshift $z_{form} = 1.87^{+>10}_{-0.58}$, is also in agreement

with the measurements of Jørgensen et al. (2014), which examined the fundamental plane of a $z = 1.27$ cluster, and implies that these galaxies are dominated by old stars which formed ~ 10 Gyr ago with some dispersion. Thus, we are seeing galaxies which are already significantly evolved when we look at MACSJ0717, consistent with a picture in which massive ETGs form their stars early and then grow passively and dissipationlessly, e.g. by minor mergers and accretion.

4.5.2 Accelerated growth?

The extremely small amount of structural evolution that we find to have taken place indicates that the galaxies in MACSJ0717 must have undergone the majority of their structural changes at earlier times. This may be a result of the very dense environment in which they are residing: indeed, other studies of galaxies in rich clusters out to $z \sim 1$ have also found no evidence for significant size or velocity dispersion evolution (Stott et al., 2011; Jørgensen & Chiboucas, 2013; Jørgensen et al., 2014; Saracco et al., 2014). Moreover, when we compare these results with those from similar studies focussing on galaxies in lower-density clusters, in which stronger structural evolution is found (e.g. Saglia et al., 2010) – and further, with those from studies of field ellipticals, which show evidence for yet stronger evolutionary trends (e.g. van der Wel et al., 2008; van der Wel et al., 2014) – a tentative picture emerges of an environment-dependent growth timescale, with galaxies in denser environments reaching their present-day sizes at earlier epochs than those in lower-density environments. Whilst we cannot comment quantitatively on this hypothesis based on the data in this study, we note that this would also be in line with the majority of studies that have directly compared the sizes of passive galaxies in high- and low- density environments (e.g. Lani et al., 2013; Delaye et al., 2014) and found the galaxies in higher-density environments to be up to 50% bigger (though see also Newman et al., 2014).

If it is indeed the case that the growth of these galaxies has been accelerated by the dense cluster environment, this would also be strong evidence in favour of merger-driven growth – which is likely to be enhanced in clusters – as opposed to growth by internal processes such as adiabatic expansion due to quasar outflows (e.g. Fan et al., 2010). It would therefore be interesting to take a deeper census of the MACSJ0717 cluster in order to establish whether the size of the galaxy population is consistent with merger rates which can account for the rate of evolution that we have observed. Of course, it is possible that our spectroscopic galaxy sample in MACSJ0717 is biased towards the largest-radius systems, in which case the cluster may also host a number of other massive but smaller

galaxies which are still undergoing some structural evolution. Whilst a deeper census would again be necessary before this could be ruled out, it is nevertheless clear that a significant population of large, massive, apparently fully evolved galaxies are already in place by $z \sim 0.5$. We also note that MACSJ0717 is an extremely massive cluster – indeed, the CLASH survey exclusively targeted strong lensing clusters – and that it may therefore be an extreme example of accelerated growth.

4.5.3 Can we trust the stellar and dynamical masses?

The stellar masses derived in Section 4.3.4 are on average higher than the dynamical masses, for the galaxies in both clusters (see Table 4.1) – implying that all the mass in these systems should be luminous. For Coma, the median ratio of stellar to dynamical mass is 1.39, while for MACSJ0717 the median ratio is 1.48. At face value this is unphysical. However, recall that our calculation of the dynamical mass does not correspond to the total dynamical mass, but is really twice the dynamical mass within one effective radius. Furthermore, with $\beta = 5$, it calculates the dynamical mass for a specific mass profile. In fact, variations in β may be as large as a factor of two for typical mass profiles of ETGs. Furthermore, when calculating the mass-to-light ratios from the BC03 stellar populations, we adopted a Salpeter IMF. Mass-to-light ratios in the r band for old, solar-metallicity SSPs are typically in the ratio 2:3 for Chabrier : Salpeter IMFs. Thus had we adopted a Chabrier (2003) or Kroupa (2001) IMF, the stellar masses would be roughly equal to the dynamical masses. We further note that we are not the first to identify stellar masses larger than dynamical masses: Peralta de Arriba et al. (2014) attribute an evolution in the stellar-to-dynamical mass being due to an evolving non-homology due to size evolution.

One caveat with our analysis is that we have attributed all the evolution to an evolution of the luminous matter as opposed to the dark matter, and the discrepancies between M_{dyn} and M_\star mean that we are unable to estimate the dark matter fractions in these galaxies and so obtain a measure of how important this assumption might be. However, more detailed studies of individual galaxies have shown that the dark matter content of ETGs only dominates at large projected radii $R > R_e$ (see e.g. Chapter 3), and so should not significantly affect stellar velocity dispersions that are measured in the central regions. It is therefore unlikely to be a significant problem in this study.

4.5.4 Can we compare MACSJ0717 with Coma?

The assumption at the foundation of this work is that the galaxies in MACSJ0717 represent an earlier evolutionary stage of the Coma galaxies: this allows us to compare their stellar

populations and so make statements about the ages of their stars and the timescales of their formation. If this assumption is not valid, it could lead to systematic errors in our age calculations, so it is important to examine it closely.

One possible problem could be the differing masses of the two clusters, with the X-ray luminosity of MACSJ0717 being more than three times greater than that of Coma – compare $L_{MACS} = 24.6 \times 10^{44} \text{ ergs}^{-1}$ (Ebeling et al., 2007) with $L_{Coma} = 7.21 \times 10^{44} \text{ ergs}^{-1}$ (Ebeling et al., 1996). In the hierarchical paradigm, more massive dark matter haloes like that of MACSJ0717 are expected to collapse earlier and so have older stars. Further, if growth is accelerated in higher-mass, higher-density systems as we have suggested, this could also lead to inconsistencies in our framework. However, the evolution that we infer is sufficiently small that even an underlying age difference of $\sim 1 \text{ Gyr}$ would not make it significant. It is therefore extremely unlikely that either of these effects would bias our inference on scales larger than our uncertainties. We also note that the good agreement between our different age measurements suggests that the framework we have set up is consistent.

4.6 Conclusions

We have constructed the colour-magnitude relation, the fundamental plane and the $M_{dyn}/L - M_{dyn}$, $M_{\star} - \sigma$ and $M_{\star} - R_e$ relations for galaxies in the cluster MACSJ0717, at $z \sim 0.5$, using archived data from the CLASH and Gemini databases, and for Coma using existing datasets. By analysing these evolution between these relations, we have reached the following conclusions.

1. The galaxies fall on an fundamental plane and an $M_{dyn}/L - M_{dyn}$ relation which are offset relative to those of Coma. The luminosity evolution implied by these offsets is $\Delta m \sim 0.6 \text{ mag}$, corresponding to a star formation epoch of $\sim 10 \text{ Gyr}$ followed by passive fading.
2. The galaxies fall on $M_{\star} - \sigma$ and $M_{\star} - R_e$ relations which are only marginally offset from those of Coma. The structural evolution implied by this is minimal, with $R_e(z) \sim (1+z)^{-0.40 \pm 0.32}$ and $\sigma(z) \sim (1+z)^{0.09 \pm 0.27}$, corresponding to galaxies which have undergone the majority of their evolution at earlier times.
3. The fundamental plane and $M_{dyn}/L - M_{dyn}$, $M_{\star} - \sigma$ and $M_{\star} - R_e$ relations, modelled together, confirm these results. Importantly, the fact that all four relations can be modelled simultaneously and consistently implies that degeneracies between the

physical parameters are not significant and that the physical scenario we have established, with evolution in luminosity, size and velocity dispersion, is consistent with and can fully account for the data. The fact that the inference from the independent colour-magnitude relation – which is also based on different assumptions about the star formation histories of the galaxies – is also consistent with these results further underlines this conclusion.

4. The small amount of structural evolution that we find in these galaxies is consistent with other studies of size evolution in cluster galaxies, but seems to be in tension with that found in studies of field ellipticals. This suggests that growth may be accelerated in high-density environments, where the rate of merging may be increased. If so, this is strong evidence that dry merging is a dominant channel of growth in these systems.
5. Taken together, these results lead to a very clear picture in which these $z \sim 0.5$ galaxies have already experienced most of their star formation and structural evolution at earlier stages in their lives.

RED NUGGETS GROW INSIDE-OUT: EVIDENCE FROM GRAVITATIONAL LENSING

Abstract

We present a new sample of strong gravitational lens systems where both the foreground lenses and background sources are early-type galaxies. Using imaging from HST/ACS and Keck/NIRC2, we model the surface brightness distributions and show that the sources form a distinct population of massive, compact galaxies at redshifts $0.4 \lesssim z \lesssim 0.7$, lying systematically below the size-mass relation of the global elliptical galaxy population at those redshifts. These may therefore represent relics of high-redshift red nuggets or their partly-evolved descendants. We exploit the magnifying effect of lensing to investigate the structural properties, stellar masses and stellar populations of these objects with a view to understanding their evolution. We model these objects parametrically and find that they generally require two Sérsic components to properly describe their light profiles, with one more spheroidal component alongside a more envelope-like component, which is slightly more extended though still compact. This is consistent with the hypothesis of the inside-out growth of these objects via minor mergers. We also find that the sources can be characterised by red-to-blue colour gradients as a function of radius which are stronger at low redshift – indicative of ongoing accretion – but that their environments generally appear consistent with that of the general elliptical galaxy population, contrary to recent suggestions that these objects are predominantly associated with clusters.

5.1 Introduction

The discovery that massive, quiescent galaxies at redshifts $z > 2$ are extremely compact (Daddi et al., 2005; Trujillo et al., 2006; van Dokkum et al., 2008; Damjanov et al., 2009, 2011) relative to their local counterparts has opened the door to important tests of our models of galaxy evolution. Whilst the hierarchical paradigm allows for the growth of passive galaxies via dissipationless mergers at a rate which may be able to account for the evolution that is required at $z \lesssim 1.5$ (e.g. Nipoti et al., 2012; Newman et al., 2012; Posti et al., 2014, but see also Sonnenfeld et al., 2014), this cannot explain the amount of evolution observed at higher redshifts or the tightness of galaxy scaling relations (Shankar et al., 2013). Adiabatic processes, such as expansion triggered by quasar feedback (Fan et al., 2010), may also be important, and the role of progenitor bias, as opposed to the growth of individual systems, remains unclear (Newman et al., 2012; Carollo et al., 2013; Belli et al., 2014).

One potentially powerful way of distinguishing between these scenarios is to quantify the morphological evolution of these galaxies. Mergers and adiabatic expansion should each leave particular imprints on the structure and stellar populations of a galaxy (Hopkins et al., 2009; Fan et al., 2010; Hilz et al., 2013), and so it should be possible to set some constraints on their relative importance in individual systems at lower redshifts. The studies of Stockton et al. (2014) and Hsu et al. (2014) attempted this at redshifts $z \sim 0.5$, using adaptive optics (AO) imaging of small galaxy samples, and found a large fraction of flattened galaxies, suggestive of disk or prolate structures, and low Sérsic indices, possibly consistent with the existence of accreted envelopes. However, discrepancies between stellar and dynamical masses in both studies (which could be indicative of high stellar velocity anisotropies resulting from their flattened morphologies) highlight the fact that their observations are really pushing the capabilities of our current observing facilities.

Strong gravitational lensing represents a way to overcome these limitations as it allows massive galaxies in the Universe to act as natural telescopes. Because lensing conserves surface brightness, a lensed background source galaxy appears not only larger, but also brighter, and this makes it possible to probe the light distributions of very small objects with high signal-to-noise data (e.g. Newton et al., 2011). Furthermore, the magnification bias of strong lensing tends to favour compact sources, making it an ideal tool to study a population of intermediate-redshift massive, compact galaxies at much higher resolutions than would otherwise be possible.

In this study, we present a new sample of thirteen early-type/early-type lens systems (EELs). These were identified as lens candidates using the SDSS spectroscopic database by

searching for spectra that could be decomposed into two ETG spectra at different redshifts, and confirmed using AO imaging in the K' -band as part of the Strong lensing at High Angular Resolution Programme (SHARP; Lagattuta et al., 2010). These now form roughly half of the SHARP sample. The first EEL has already been shown to be a massive, compact ETG at redshift $z = 0.63$, and was found to require a two-component Sérsic model to accurately fit the surface brightness profile, including an extended low-surface-brightness component (Auger et al., 2011), in line with expectations of the effect of merging and accretion on high-redshift nuggets (Hopkins et al., 2009). However, those models were based on single-band AO imaging with an uncertain PSF (whose broad wings generally affect the measurement of the low-surface brightness outskirts); we now have HST/ACS images for all of the EELs, facilitating a much more thorough study. Here, we analyse the entire sample to investigate and exploit the idea that this relatively unexplored class of gravitational lenses naturally selects compact nugget descendants.

This chapter is structured as follows: we present the data in Section 5.2 and our lens modelling methods and results in Sections 5.3 and 5.4. We then investigate and discuss the properties of the source galaxies in Sections 5.5 and 5.6 and finally conclude in Section 5.7.

5.2 Data

As summarised by Auger et al. (2011), EEL candidates were identified by searching the SDSS spectroscopic database for spectra that could be decomposed into two ETG spectra at different redshifts (similarly to the method employed by SLACS; Bolton et al., 2006, though that survey searched for emission lines in the background sources). SDSS imaging was used to reject lens candidates that were clearly resolved into two galaxies, and a probability for lensing was determined based on the velocity dispersion of the foreground galaxy. Fourteen candidates were observed in the K' -band using NIRC2 with laser guide star adaptive optics (LGS-AO) on Keck II over a range of dates from August 2009 until May 2012, most as part of SHARP, and all were confirmed as lenses. The data were reduced as described by Auger et al. (2011), with images taken using the wide camera drizzled to a scale $0.03''/\text{pixel}$ and those taken using the narrow camera drizzled to a scale of $0.01''/\text{pixel}$. The zeropoints for these data were calibrated against 2MASS photometry, which includes robust detections of all of the systems except J0913 and J1446. For these two objects, we used 2MASS photometry to determine the zeropoints for other targets observed on the same nights, finding negligible scatter throughout the nights.

These EELs were also observed using HST/ACS as part of the programme GO 13661 (PI: Auger). Two dithered exposures of duration ~ 500 s were observed in the I -band

EEL	RA (deg)	DEC (deg)	z_l	z_s
J0837	08:37:01.21	+08:01:17.89	0.4248	0.6406
J0901	09:01:21.25	+20:27:40.41	0.3108	0.5860
J0913	09:13:45.65	+42:37:30.81	0.3946	0.5390
J1125	11:25:13.89	+30:58:05.59	0.4419	0.6884
J1144	11:44:28.40	+15:40:39.36	0.3715	0.7050
J1218	12:18:06.67	+56:48:05.12	0.3177	0.6000
J1248	12:48:47.82	+47:11:05.81	0.3042	0.5276
J1323	13:23:59.07	+39:46:33.24	0.3192	0.4637
J1347	13:47:04.96	−01:01:03.57	0.3974	0.6289
J1446	14:46:30.20	+38:56:56.41	0.3175	0.5858
J1605	16:05:23.28	+38:11:53.95	0.3065	0.5418
J1606	16:06:07.09	+22:35:11.35	0.3810	0.6545
J1619	16:19:12.63	+20:24:27.97	0.3635	0.6132
J2228	22:28:40.80	−00:18:16.84	0.2387	0.4366

Table 5.1: Positions and redshifts, for both source and lens, of the fourteen EELs.

(F814W), and another set of two dithered exposures of ~ 500 s were obtained in the V-band (F555W for sources at redshift $z < 0.55$ or F606W for $z > 0.55$, in order to straddle the 4000Å break). The ACS data were reduced using ASTRODRIZZLE and were drizzled to a scale of $0.05''/\text{pixel}$. There are a small number of artefacts in the resulting images due to the limited number of exposures in each band, and these are masked in the subsequent analysis. The positions on the sky of these fourteen systems are summarised in Table 5.1, along with the redshifts of both source and lens.

5.3 Lens modelling

One of the main aims of this study is to robustly measure the sizes, morphologies and masses of the source galaxies in order to compare their size-mass relation both with other galaxies at similar redshifts and with high-redshift nuggets; we therefore choose to model their light distributions using elliptical Sérsic profiles. An alternative would be to make pixellated source reconstructions (e.g. Warren & Dye, 2003; Vegetti & Koopmans, 2009) from which half-light radii could be measured. However, this would add an extra level of uncertainty to the final size and magnitude measurements and complicate the interpretation of the sizes; nevertheless, for a small number of systems, we do carry out inference based on pixellated sources as a verification of our parametric lens models, but we do not use these in the analysis. (We also make pixellated reconstructions of all the EELs

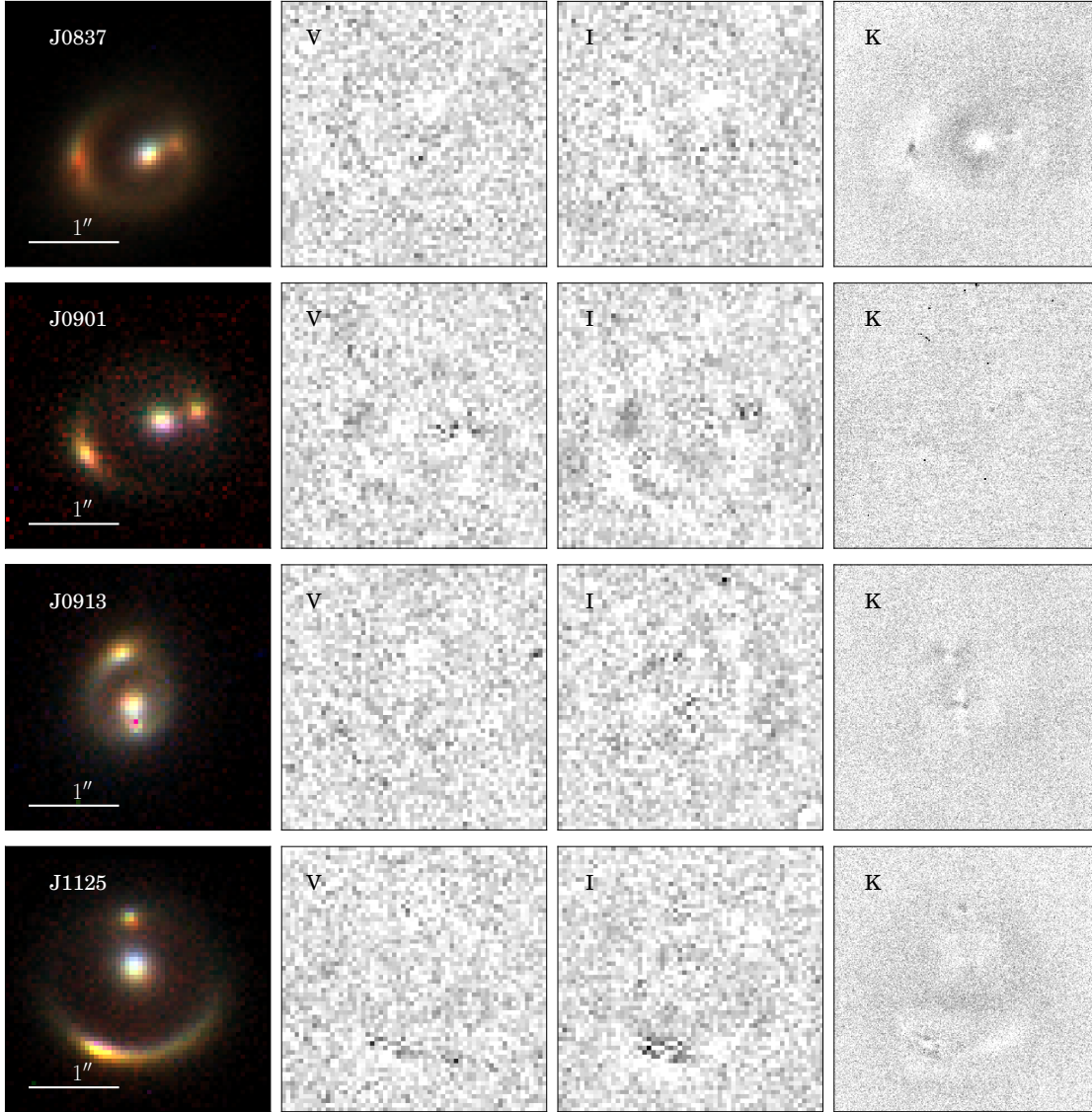


Figure 5.1: From left to right, we show the colour image combining all three bands of data and the residuals for the V , I and K' bands, for the best model (i.e., 1C/2C) for each system as given in Table 5.2. All cutouts are 3 arcseconds on a side.

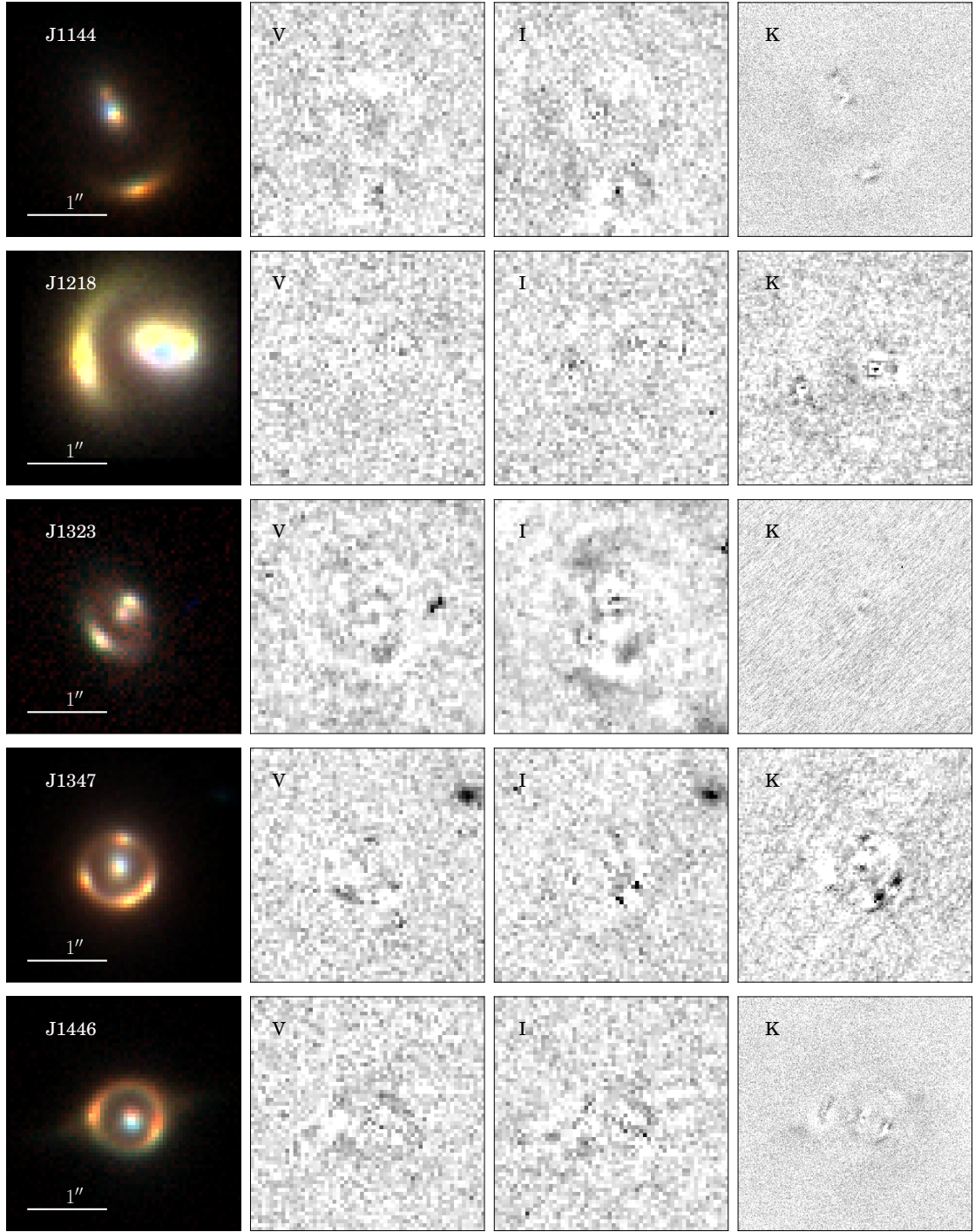


Figure 5.1: continued.

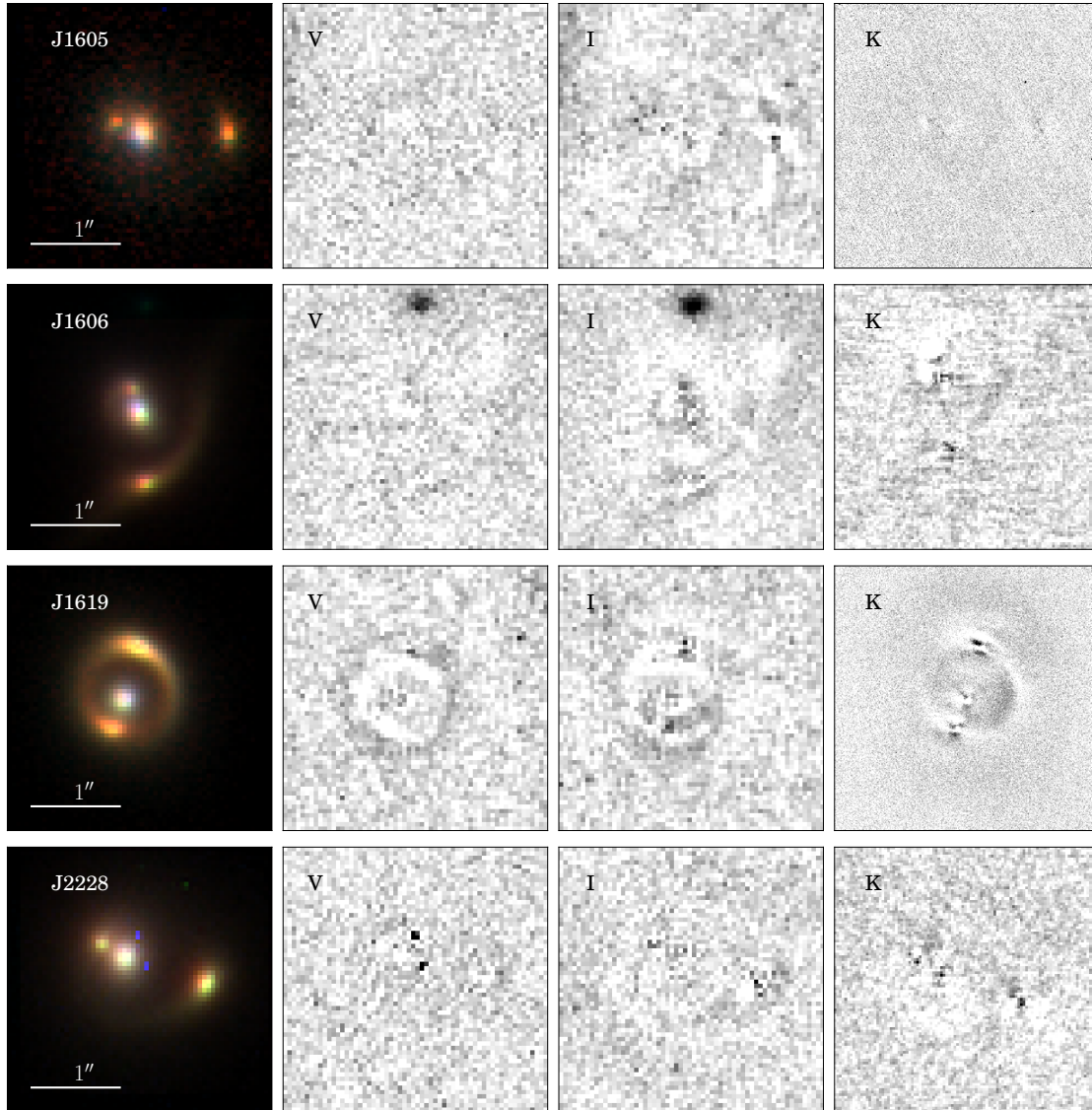


Figure 5.1: continued.

sources, and show these in Appendix D.) Further, single-component Sérsic profiles are a standard way of modelling surface brightness distributions for both lensed and unlensed galaxies at all redshifts (e.g. Shen et al., 2003; Newton et al., 2011; van der Wel et al., 2014), so modelling our lensed sources in a similar way allows a straightforward comparison with other studies (see Marshall et al., 2007, for a discussion of the advantages and limitations of parametric source modelling).

Equally, some sources with more complex light distributions may not be well described by single-component Sérsic models – for instance, those containing bars or bulges and disks – and, from a lensing point of view, it is important to verify that any residuals in the model are a result of the shortcomings of the light profile that has been imposed, rather than the mass model. Further, it is important to be able to measure the total flux from the source and assess any uncertainty or bias introduced by assuming a single Sérsic profile. For each system, we therefore create two ‘best’ models, the first using a single Sérsic component for the source (which we call a 1C or ‘one-component’ model) and the second with two Sérsic components (which we call a 2C or ‘two-component’ model); for some systems, the 1C model allows us to describe the data down to the noise level, and we do not create 2C models in these cases. For the foreground galaxy, we also use either one or two components. In all models with more than one component for either the foreground galaxy or the source, we require the two components to be concentric, but allow their position angles and ellipticities to be independent.

For each first Sérsic component, we therefore have six free non-linear parameters – $(x, y, q_1, \phi_1, R_{e,1}, n_1)$ – where (x, y) gives the centroid, q_1, ϕ_1 describe the axis ratio and position angle and $R_{e,1}, n_1$ describe the half-light radius and index of the Sérsic profile. For each second Sérsic component, we have four free parameters: $(q_2, \phi_2, R_{e,2}, n_2)$. We model the lensing mass of the foreground galaxy using an elliptical power law distribution (calculating deflection angles according to the prescription of Barkana, 1998) and allow for an external shear; whilst the simpler, more common singular isothermal ellipsoid (SIE) distribution has been shown to provide a good approximation to the lens potential on galaxy scales (e.g. Treu & Koopmans, 2004), our focus is on measuring reliable and robust sizes and we therefore want to eliminate as much potential bias in our source models as possible. Our mass model therefore has eight free parameters – $(x_l, y_l, q_l, \phi_l, R_{Ein}, \eta, \gamma_{ext}, \phi_{ext})$ – where (x_l, y_l) describe the centroid of the mass, q_l, ϕ_l give its axis ratio and position angle, R_{Ein}, η give the Einstein radius and the power law index of the 3D density profile $\rho \propto r^{-(\eta+1)}$ and γ_{ext}, ϕ_{ext} give the magnitude and position angle of an external shear. We do not require the mass and light of the lens galaxy to be concentric or aligned.

For a given set of these *non-linear* parameters, we determine the *linear* amplitude

of each surface brightness component by evaluating the foreground galaxy profile in the image plane and the source galaxy profile in the source plane, given the deflection angles of the mass model. We do not subtract the foreground galaxy light prior to the modelling due to the covariance between the foreground and background light. These are especially covariant in the EELs as compared to other lens systems due to their generally small Einstein radii and similar colours, which result in a very large amount of overlap between the source and lens light.

The model is then convolved with the PSF; for the HST images, we use a nearby unsaturated star for the PSF in each band, whereas for the K' band data, with an unstable PSF and often with no reference star in the field of view, we model the PSF as the sum of three (concentric but not aligned) elliptical Gaussian profiles, and infer the properties of these Gaussians along with the other model parameters. We then use a non-negative least squares linear inversion to find the best combination of the foreground lens and background source light components and a uniform background component, and thereby calculate the likelihood for the data \vec{D} , given the non-linear parameters of the model \vec{M} , as

$$(5.1) \quad \ln L(\vec{D}|\vec{M}) = -\frac{1}{2} \sum_i \left(\frac{d_i - m_i}{\sigma_i} \right)^2$$

where d_i, m_i, σ_i are the i^{th} pixel in the data image, model image and noise map respectively, and the sum is over all unmasked pixels (for some systems, bright interloping objects must be masked by hand). Given uniform priors on all the non-linear parameters, we can then infer the posterior distribution, $p(\vec{M}|\vec{D})$, of the model given the data in a Bayesian way using an MCMC exploration. To ensure that the parameter space is fully explored when the posterior is not necessarily uni-modal, we use the parallel-tempered version of EMCEE with three temperatures.

We begin by modelling the HST V and I bands jointly, requiring the light and mass profiles to be the same in both filters (accounting for their different PSFs and spatial shifts between bands), and allowing each surface brightness component to contribute a different amount to the flux in each filter. We then model the K' band separately, fixing the lensing galaxy's mass and light profiles to those inferred from the HST data and inferring the PSF and the source profile. The impetus for remodelling the source in the K' band, but not the foreground galaxy, is that we are particularly interested in the structure of the potentially nugget-like source galaxies here, including the possibility that they might exhibit strong colour gradients due to ongoing or recent evolution, which would lead to smaller measured sizes in the K' -band. We test this rationale by creating models for a subset of the EELs in which we also fix the source profile, and infer just the PSF, and find that the residuals

are considerably worse in a number of cases. We additionally create models in which the foreground galaxy light profile is also allowed to change (though the mass remains fixed), and recover a posterior distribution that is consistent with the HST models. In our analysis of the size-mass relation, we opt to use the sizes from the HST data, as these are generally more robust since they are not dependent on any inference on the PSF.

We also create models in which the three bands are fitted simultaneously. In this case, we infer the lens mass and light profiles, which are the same in all filters, as well as the K' band PSF and the source profile, where the latter is now a single Sérsic component with a wavelength-dependent effective radius given by

$$(5.2) \quad \log(R_e/\text{arcsec}) = \alpha_R \log(\lambda/6000\text{\AA}) + \beta_R$$

for wavelength λ . This model therefore allows for colour gradients while modelling all three bands in a consistent way, and provides an important consistency check for our inferred mass profiles. It is also informative as a further way of distinguishing between different red nugget growth scenarios (e.g. Fan et al., 2010; Wuyts et al., 2010; Hilz et al., 2013; Ishibashi et al., 2013) which make distinct predictions for the extent and colours of the stellar populations that should be observed. These models are treated separately in Section 5.6.3.

EEL	z_l	z_s	R_{Ein} (arcsec)	η	q_{lens}	ϕ_{lens} (deg)	γ_{ext}	ϕ_{ext} (deg)	N
J0837	0.4248	0.6406	0.56 ± 0.01	1.20 ± 0.01	0.76 ± 0.01	29.80 ± 1.18	0.06 ± 0.01	-116.96 ± 0.59	1C
J0901	0.3108	0.5860	0.67 ± 0.01	1.07 ± 0.01	0.82 ± 0.01	5.51 ± 1.29	-0.04 ± 0.01	12.35 ± 1.88	1C
J0913	0.3946	0.5390	0.42 ± 0.01	1.24 ± 0.02	0.79 ± 0.02	-121.25 ± 2.24	0.04 ± 0.01	-35.66 ± 3.45	2C
J1125	0.4419	0.6884	0.86 ± 0.01	0.96 ± 0.02	0.92 ± 0.01	112.54 ± 1.94	0.08 ± 0.01	97.52 ± 0.52	2C
J1144	0.3715	0.7050	0.68 ± 0.01	1.08 ± 0.02	0.75 ± 0.02	-57.30 ± 0.90	-0.04 ± 0.01	30.15 ± 3.16	2C
J1218	0.3177	0.6000	0.68 ± 0.01	1.11 ± 0.01	0.81 ± 0.01	-37.48 ± 2.06	-0.02 ± 0.01	-87.30 ± 4.68	1C
J1323	0.3192	0.4637	0.31 ± 0.01	1.01 ± 0.01	0.76 ± 0.01	-66.18 ± 2.14	0.02 ± 0.01	11.25 ± 8.06	1C
J1347	0.3974	0.6289	0.43 ± 0.01	1.23 ± 0.00	0.62 ± 0.00	90.29 ± 0.38	-0.01 ± 0.01	-64.55 ± 4.40	2C
J1446	0.3175	0.5858	0.41 ± 0.01	1.44 ± 0.02	0.79 ± 0.01	-73.50 ± 1.31	0.01 ± 0.01	66.29 ± 7.27	2C
J1605	0.3065	0.5418	0.64 ± 0.01	1.25 ± 0.02	0.67 ± 0.02	98.76 ± 1.66	0.06 ± 0.01	-26.92 ± 3.69	2C
J1606	0.3810	0.6545	0.52 ± 0.01	1.21 ± 0.01	0.62 ± 0.01	-53.91 ± 0.96	0.09 ± 0.01	25.81 ± 1.73	2C
J1619	0.3635	0.6132	0.50 ± 0.01	1.28 ± 0.03	0.97 ± 0.02	-68.87 ± 16.58	-0.06 ± 0.01	-34.39 ± 2.21	2C
J2228	0.2387	0.4366	0.60 ± 0.01	1.12 ± 0.01	0.96 ± 0.01	-59.87 ± 7.59	-0.06 ± 0.01	1.59 ± 2.01	2C

Table 5.2: A summary of the lens models, inferred using the HST V- and I-band data, with statistical uncertainties. We present the lens and source redshifts (measured from the SDSS spectra), the Einstein radius in arcsec, the power-law index η , the ellipticity and position angle of the lens and the magnitude and position angle of the extrnal shear. The final column denotes the ‘best’ model for each system, which is either 1C (one Sérsic component) or 2C (two Sérsic components).

5.4 Modelling Results

The results of our lens modelling are summarised in Tables 5.2 (lens mass models) and 5.3 (source light models). We omit the lens J1248 because the lensing galaxy is clearly an edge-on disk galaxy and we find that the elliptical power law plus external shear mass model does not adequately describe the lensing potential. For the sources, we present the *VIK* magnitudes, the effective radii R_e , and Sérsic indices for both 1C and 2C models, and the axis ratio q for the 1C model. Since we are primarily interested in the source properties in this study, we do not include the inference on the foreground galaxy light distributions here; these will be presented in a future work. We then present the images, models and signal-to-noise residuals for each EEL in the three bands in Figure 5.1.

Whilst our focus is to create reproducible 1C models which are easy to interpret and compare with other studies, a number of systems presented peculiar features during the modelling process which required small changes to the main model, or simply offered interesting insights into the systems. These are summarised in Appendix D. For a number of these, we also created pixellated models of the source, using techniques similar to those described in Vegetti & Koopmans (2009), subtracting our best parametric model for the foreground galaxy and inferring the lensing mass distribution and regularisation. Where appropriate, these are also explained in Appendix D. The pixellated models for all systems are presented in Appendix C.

5.4.1 Accurately modelling the EELs

In some cases, the reason for the failure of the 1C model is readily apparent. J1606, for instance, is dominated by a disk but also has a very prominent bulge which the single-component model simply cannot reproduce, and the same is true for J1446’s disk; more generally, we point out that the one-component models tend to fail where the surface brightness profile is particularly extended or has a low-surface-brightness envelope, in which case the Sérsic index becomes large in an attempt to describe both the bright, compact central structure and the extended brightness at larger radii. This raises an important point: the surface brightness structures of galaxies are generally much more complex than single Sérsic profiles, and the fact that our sources are lensed and therefore imaged with excellent resolution, given their redshifts, means that we cannot get away with overly simple models here. We test the degree of complexity that seems to be required by adding third components to our models, and find that these tend to be poorly constrained and associated with very small amounts of flux. It seems, then, that double Sérsic profiles are adequate – and usually necessary – to describe a typical EEL source.

	2C			1C			2C			
EEL	m_V (mag)	m_I (mag)	$m_{K'}$ (mag)	R_e (kpc)	n	q	R_e (kpc)	n_{env}	n_{bulge}	B/T_I
J0837	21.31 ± 0.02	19.63 ± 0.02	18.07 ± 0.03	4.42 ± 0.27	4.73 ± 0.19	0.50 ± 0.01	—	—	—	—
J0901	22.08 ± 0.02	20.48 ± 0.02	19.52 ± 0.03	3.26 ± 0.19	5.11 ± 0.14	0.72 ± 0.01	—	—	—	—
J0913	22.12 ± 0.02	19.97 ± 0.02	18.21 ± 0.03	4.68 ± 0.29	4.83 ± 0.13	0.55 ± 0.01	4.11 ± 0.17	3.13 ± 0.34	6.78 ± 1.23	0.72 ± 0.05
J1125	23.41 ± 0.02	21.85 ± 0.02	19.83 ± 0.03	4.32 ± 0.46	6.24 ± 0.29	0.71 ± 0.01	1.17 ± 0.02	0.92 ± 0.06	3.06 ± 0.46	0.71 ± 0.06
J1144	21.19 ± 0.02	19.77 ± 0.02	19.01 ± 0.03	8.54 ± 0.68	6.85 ± 0.19	0.83 ± 0.02	9.64 ± 0.28	0.94 ± 0.07	4.08 ± 0.19	0.61 ± 0.06
J1218	21.12 ± 0.02	19.59 ± 0.02	17.89 ± 0.03	6.79 ± 0.33	4.66 ± 0.09	0.66 ± 0.01	—	—	—	—
J1323	21.83 ± 0.02	19.96 ± 0.02	17.35 ± 0.03	1.82 ± 0.11	4.97 ± 0.22	0.51 ± 0.01	—	—	—	—
J1347	22.27 ± 0.02	20.91 ± 0.02	19.74 ± 0.03	3.96 ± 0.33	8.51 ± 0.34	0.89 ± 0.02	5.39 ± 0.49	1.29 ± 0.19	8.09 ± 0.40	0.40 ± 0.05
J1446	22.23 ± 0.02	20.71 ± 0.02	18.96 ± 0.03	2.50 ± 0.09	4.13 ± 0.09	0.53 ± 0.01	1.59 ± 0.04	0.50 ± 0.02	3.98 ± 0.23	0.47 ± 0.07
J1605	22.62 ± 0.02	20.44 ± 0.02	18.38 ± 0.03	3.36 ± 0.13	4.16 ± 0.09	0.71 ± 0.01	2.56 ± 0.05	1.18 ± 0.08	2.73 ± 0.31	0.72 ± 0.06
J1606	21.57 ± 0.02	19.93 ± 0.02	17.91 ± 0.03	15.91 ± 0.42	8.40 ± 0.11	0.24 ± 0.00	3.12 ± 0.12	0.53 ± 0.01	7.74 ± 0.28	0.26 ± 0.04
J1619	21.17 ± 0.02	19.64 ± 0.02	18.51 ± 0.03	7.32 ± 0.73	6.17 ± 0.23	0.69 ± 0.01	5.24 ± 0.20	1.49 ± 0.15	5.07 ± 0.35	0.44 ± 0.08
J2228	21.27 ± 0.02	19.60 ± 0.02	18.61 ± 0.03	12.32 ± 0.77	9.41 ± 0.19	0.80 ± 0.01	4.15 ± 0.08	0.66 ± 0.03	4.65 ± 0.23	0.52 ± 0.05

Table 5.3: A summary of source galaxy properties, with statistical uncertainties. Columns 2-4 give the unlensed m_V , m_I and $m_{K'}$ apparent magnitudes, calculated for the ‘best’ model (i.e. 1C or 2C, as given in Table 5.2). Columns 5 - 7 give the effective radius, Sérsic index and axis ratio for the one-component models. Columns 8 - 11 give the corresponding properties of the two-component models (where they exist): here, the effective radius is that containing half the total (summed) light, taking into account both components. n_{env} and n_{bulge} are the Sérsic indices of the envelope-like and bulge-like components and B/T_I is the bulge-to-total ratio measured in the I -band.

EEL	$\log(M_{\star}/M_{\odot})$
J0837	11.67 ± 0.04
J0901	11.19 ± 0.04
J0913	11.30 ± 0.08
J1125	11.01 ± 0.06
J1144	11.57 ± 0.05
J1218	11.63 ± 0.05
J1323	11.21 ± 0.06
J1347	11.12 ± 0.08
J1446	11.11 ± 0.09
J1605	11.09 ± 0.09
J1606	11.48 ± 0.06
J1619	11.47 ± 0.12
J2228	11.26 ± 0.05

Table 5.4: Stellar masses for the source galaxies, inferred from the photometry using the BC03 stellar population models and assuming a Chabrier IMF.

An added complication in the modelling of these systems is that the surface brightness profiles of both foreground and background galaxy are unknown, and are presumably comparable in both colour and brightness; it is therefore possible that they are degenerate. We find, however, that this is generally not the case when both are modelled simultaneously, though it is possible that modelling in which the source is first masked and the foreground light modelled separately and then subtracted could be problematic due to the small Einstein radii of these systems.

On the other hand, we do find that the robustness of the inference on the light profiles relies on carrying out the modelling using image cutouts which capture a sufficient fraction of the light, and that this fraction is surprisingly large: our final cutout radius is \sim five times the effective radius of the largest Sérsic component in the foreground+background model (typically $\sim 5''$), and we find that modelling the same system on smaller cutouts leads to systematically different inference on the Sérsic indices, with a larger number of foreground galaxies having components with $n < 1$, and the source galaxies having systematically larger n . Both of these cases increase the amount of light at large radii, beyond the extent of the cutout, where it cannot be penalised by data. This emphasises the necessity of modelling the full region surrounding the lens system, in spite of the small Einstein radii of the EELs.

5.5 Source galaxy properties

The combination of high-resolution imaging with the magnification due to lensing means that the EELs sources can be resolved in great detail. In this section, we present inference on their stellar masses and their size-mass relation, and point towards some characteristic features in their morphologies relative to those of the low-redshift SDSS galaxy population.

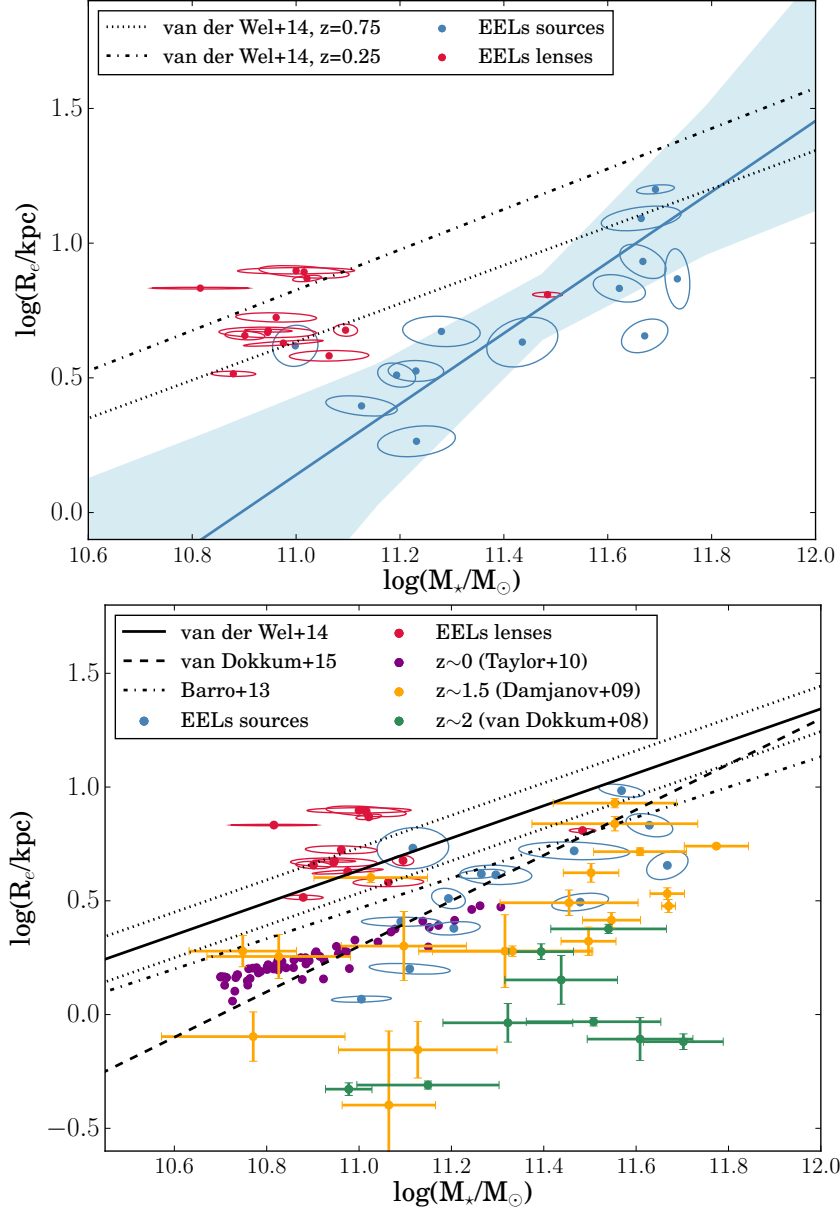


Figure 5.2: The size-mass relation for source galaxies (blue) and lens galaxies (red). Top: 1C models, with the size-mass relations for the global ETG population from van der Wel et al. (2014) plotted for reference. The size-mass relation for the source population is well below the van der Wel et al. (2014) relation across a large part of the mass range. Bottom: 2C models, with the criteria for compactness used in Barro et al. (2013) and van Dokkum et al. (2015) plotted for reference in addition to the van der Wel $z = 0.75$ relation with its intrinsic scatter. Also plotted are the red nugget populations from Taylor et al. (2010), Damjanov et al. (2009) and van Dokkum et al. (2008), which suggest an evolution towards increasing size at lower redshifts. Our source galaxies are much more consistent with this trend within the red nugget populations, whereas the lens galaxies are consistent with the global population (though they span a very small range in stellar mass).

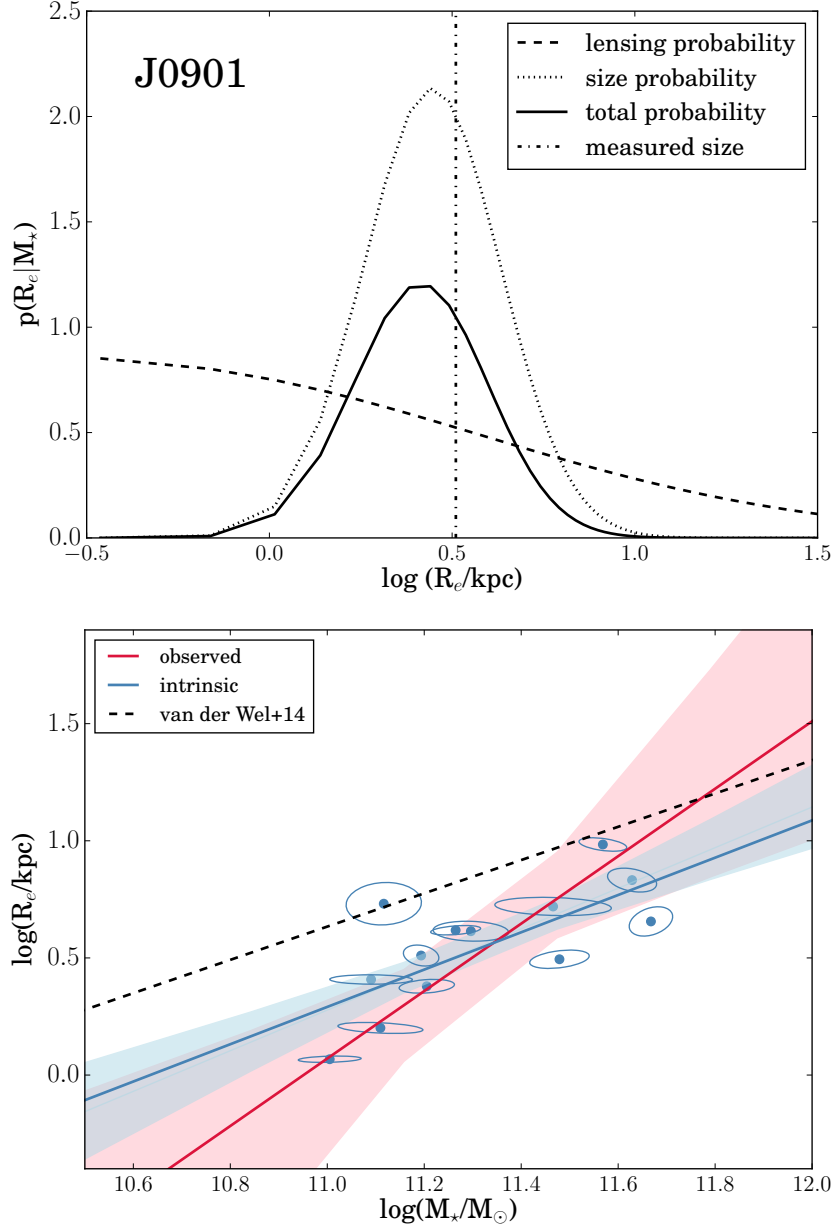


Figure 5.3: Recovering the intrinsic size-mass distribution of compact galaxies. Top: the intrinsic size-mass relation (dotted line) at a given stellar mass is modified by the bias introduced by differential magnification (dashed line) to yield the overall probability of observing an EEL source with a particular stellar mass and effective radius (solid line). Bottom: The intrinsic size-mass relation (here for 2C models) is shallower than the observed relation. Relative to the $z = 0.75$ van der Wel et al. (2014) relation, it is offset to smaller sizes but has a consistent slope.

model	α_{SM}	β_{SM}	σ_{SM}	μ_{SM}	τ_{SM}
observed relation					
1C	$0.16^{+0.27}_{-0.44}$	$1.27^{+0.90}_{-0.55}$	$0.11^{+0.11}_{-0.08}$	$11.45^{+0.08}_{-0.08}$	$0.18^{+0.10}_{-0.08}$
2C	$0.07^{+0.25}_{-0.40}$	$1.44^{+1.15}_{-0.71}$	$0.13^{+0.11}_{-0.09}$	$11.33^{+0.08}_{-0.08}$	$0.13^{+0.09}_{-0.07}$
intrinsic relation					
1C	$0.36^{+0.11}_{-0.11}$	$0.83^{+0.22}_{-0.23}$	$0.19^{+0.06}_{-0.04}$	$11.43^{+0.08}_{-0.08}$	$0.28^{+0.07}_{-0.05}$
2C	$0.28^{+0.10}_{-0.09}$	$0.87^{+0.24}_{-0.25}$	$0.18^{+0.05}_{-0.04}$	$11.32^{+0.07}_{-0.07}$	$0.24^{+0.07}_{-0.05}$

Table 5.5: Inference on the size-mass relation for the source galaxy population, for 1C and 2C models. The parameters correspond to those defined in Equations 3 and 4; we model the sources as following the linear relation $\log(R_e/kpc) = \beta_{SM} \log(M_\star/10^{11}M_\odot) + \alpha_{SM}$ with an intrinsic scatter σ_{SM} in the $\log R_e$ direction, and allowing the masses to be drawn from an underlying Gaussian distribution $p(\log M_\star) = N(\mu_{SM}, \tau_{SM}^2)$.

5.5.1 Stellar masses

As the EELs were originally identified in SDSS, each combined source+lens system also has measured *ugriz* photometry in the SDSS database, and we can use this in addition to our *VIK* photometry to make inference on the physical properties of both source and lens. We do not use their 2MASS photometry, as this gives little extra information alongside our NIRC2 photometry (which also has the advantage of giving magnitudes for lens and source separately, unlike the 2MASS and SDSS photometry, and thus helps to break the degeneracy between source and lens light). We also reject the SDSS *u*-band photometry, as it has very large uncertainties due to the lack of flux from ETGs at such blue wavelengths. Note that, for objects with *VIK* photometry based on two-component models, we infer *total* stellar masses using the total magnitudes, rather than assigning each component its own mass; this is because our Sérsic profiles are only parameterisations of the light distribution and do not necessarily represent two distinct physical components.

We then infer the stellar masses of both source and lens galaxy using the composite stellar population synthesis models of Bruzual & Charlot (BC03, 2003). Our code uses these models to compute the magnitudes, for a specified set of filters and redshift, on a grid of stellar age T , metallicity Z , dust extinction τ_v and time constant τ of an exponentially decaying star formation history, and constructs a spline interpolation model which allows magnitudes to be evaluated at arbitrary points within the grid. In this approach, we follow the methods developed by Auger et al. (2009). We then explore the posterior probability distribution of these parameters, along with the stellar masses of the two objects, by MCMC sampling, noting that, as we are combining photometry for the separated source and lens light (from HST and Keck) with photometry for the combined system (from SDSS),

the likelihood is non-linear in the logarithms of the lens and source masses M_* . We use uniform priors on T , τ , $\log \tau_v$, $\log Z$ and $\log M_*$ for each object and model the source and lens photometry simultaneously, as stated previously. As discussed by Auger et al. (2009) and Newton et al. (2011), despite large degeneracies between a number of the parameters – such as T and Z , and T and τ – the stellar masses are not significantly affected by these degeneracies and this makes it possible to constrain them with uncertainties of $\sim 0.05 - 0.1$ dex for a given IMF. We adopt a Chabrier IMF, in keeping with previous studies of the size-mass relation (e.g. Shen et al., 2003; van der Wel et al., 2014), but note that the use of a Salpeter IMF – which recent evidence suggests may be more suitable for massive ETGs (Auger et al., 2010b) – would increase the stellar masses by a factor of ~ 1.7 . The ‘best model’ (i.e. 1C/2C) stellar masses for the sources are presented in Table 5.4.

5.5.2 The observed size-mass relation

We use the half-light radii inferred from the lens modelling and the stellar masses inferred from the photometry to construct the size-mass relation for both 1C and 2C models for the EELs sources. In this section, we model the *observed* relation, ignoring the selection function of the sample; we then attempt to recover the *intrinsic* size-mass relation in the following section.

We model the observed size-mass relation of the source population as a normal distribution,

$$(5.3) \quad \log(R_e/\text{kpc}) = N(\beta_{SM} \log(M_*/10^{11} M_\odot) + \alpha_{SM}, \sigma_{SM}^2),$$

accounting for covariance between the size and stellar mass measurements, and treating the masses as being drawn from an underlying normal distribution with mean μ_{SM} and standard deviation τ_{SM} ,

$$(5.4) \quad p(\log M_*) = N(\mu_{SM}, \tau_{SM}^2).$$

This is consistent with the fact that, as a result of the EELs selection algorithm and the galaxy mass function, we do not expect the parent distribution of stellar masses $p(\log M_*)$ to be flat. In this approach we follow the formalism presented by Kelly (2007). We note that we model parent distributions using single normal distributions in what follows, but have verified that our inference is robust against increases in the number of normal distributions used.

The inferences for both 1C and 2C models are summarised in Table 5.5, and the relations are shown in Figure 5.2. For comparison, we also show the EELs foreground

lensing galaxies, though it is clear from the figure that this population lacks the dynamic range in stellar mass to allow us to identify any meaningful trends. It is interesting to note that the sources have a larger mean mass than the lenses; we find $\mu_{SM} = 11.03$ (in units of $\log(M_{\star}/M_{\odot})$) for the lens galaxies, which is 2 times smaller than the $\mu_{SM} = 11.32$ that we calculate for the 2C models of the sources. As the cross-section for strong gravitational lensing scales approximately with lensing mass, it is an expectation that the lens galaxies will form a massive population. However, large masses for the sources are not necessarily expected, and this arises here as a result of the specific selection criteria for the EELs – that is, detecting their spectra in the SDSS fibres requires that they be bright, with (at least a magnified) flux comparable to that of the lens galaxy. This underlines the fact that the EELs sources, as well as the lenses, constitute a massive population.

It can be seen from Figure 5.2 that the EELs sources are compact. We also plot the fits to the size-mass relation from van der Wel et al. (2014) – both at $z = 0.25$ and $z = 0.75$, which are chosen to bracket the redshifts of the EELs sources – in the top panel. Nearly all the sources lie distinctly below these lines. For comparison, the lens galaxy sample straddles the $z = 0.25$ size-mass relation, as might be expected given their average redshift $\bar{z}_l = 0.35$. In the top panel, we show the EELs lenses and sources alongside the red nugget populations from van Dokkum et al. (2008), Damjanov et al. (2009) and Taylor et al. (2010), which span redshifts between $z \sim 2$ and $z \sim 0$, in addition to the compactness criteria for classifying high-redshift nuggets used by Barro et al. (2013) and van Dokkum et al. (2015) and the global $z = 0.75$ size-mass relation of van der Wel et al. (2014), along with its intrinsic scatter. Seen in this context, the EELs source population appears to occupy a region closer to the red nuggets than to ‘normal’ ETGs.

We note, however, that the relations shown on this plot are meant to define some sort of boundary between ‘compact’ and ‘non-compact’ objects, with the former all lying below it; our EELs sources are instead scattered above and below these lines. Specifically, nine out of the thirteen systems would be classed as compact according to Barro et al. (2013)’s criterion, whereas van Dokkum et al. (2015)’s slightly stricter definition reduces this to seven – though, due to differences in the two criteria at high and low masses, these two subsamples do not completely overlap.

Given the distinct position in size-mass space of our sources, in addition to the diversity of conflicting compactness definitions that exist, we do not think it is valuable to classify our sources in this way. Rather, we simply note that they seem to be significantly more compact than the majority of ETGs at similar masses and redshifts, and may be better associated with the red nugget population. For instance, they may represent red nuggets at some intermediate stage of their evolution, caught in the act of accreting matter. This is

a possibility we consider in more detail in Section 5.6.

5.5.3 The intrinsic size-mass relation

The EELs sample is subject to a non-trivial selection function which steepens the slope of the size-mass relation that is observed. We now model this to recover the intrinsic size-mass relation.

The selection function of the EELs sources is threefold. Firstly, the source must be lensed by the foreground object; this relates to the cross-section for lensing. Secondly, the inclusion of an EEL in the SDSS spectroscopic sample requires the lens+source system as a whole to fulfil the criteria of the SDSS target selection algorithm (Strauss et al., 2002), which itself is non-trivial, though the main effect is that the system is bright. Finally, the EEL must pass our spectroscopic search, which is somewhat subjective but imposes criteria such as the lensed source flux being comparable to the lens galaxy flux and the redshifts of the two objects approximately satisfying $0.1 \lesssim z \lesssim 0.7$. The combination of these different conditions leads to some selection function which modifies the intrinsic population of compact galaxies to the population of EELs sources that we observe.

Of these three contributions, the latter two are difficult to quantify and should not introduce any large bias into our measurement of the size-mass relation, although they will push us to the high-mass end of the relation. On the other hand, the first – the lensing cross section – introduces a selection function such that we are relatively more efficient at selecting compact galaxies at lower masses. We can understand this as follows: differential magnification introduces a bias towards smaller objects (closer to the line-of-sight of the lens), whereas, for a given size, there is no bias as a function of luminosity, and therefore stellar mass (above a limit set by the latter two criteria discussed above; note also that this is not in contradiction with the well-known magnification bias, which encodes the fact that the number density of sources increases with decreasing brightness, and not that the probability of a single object being lensed increases with decreasing brightness). The result of this is that an object of fixed luminosity becomes increasingly likely to be seen in the lensed population relative to the intrinsic population as it becomes more compact.

This effect is demonstrated by the dashed curve in the top panel of Figure 5.3, which shows the magnification (which we use as the lensing probability) for the EEL J0901 as a function of the effective radius of the source. The shape of the curve shows that the bias is towards smaller sizes (and therefore lower-mass objects). Of course, the probability of this lensing occurring in the physical Universe also depends on the *intrinsic* distribution of stellar mass and size, i.e. the intrinsic distribution of compactness, which, given the stellar

mass of an object, gives the probability distribution of that object having a particular effective radius and which is what we ultimately would like to infer. In the figure, our final inference on this distribution (i.e. the intrinsic size-mass relation, see below, evaluated at the stellar mass of J0901) is shown by the dotted curve, and the corresponding probability distribution of effective radii for the EEL, given that it has been observed (i.e. the *observed* size-mass relation, evaluated at the stellar mass of the EEL) is shown by the solid black curve. Thus the intrinsic size-mass distribution is modified by the bias introduced by lensing due to differential magnification.

We use this setup to infer the underlying size-mass relation, given the size-mass relation that we observe. We do this using an MCMC exploration, positing an underlying size-mass relation as in Equation 5.3, and using this to calculate the probability that each EEL would be observed as a function of radius. This gives a likelihood function for the i^{th} EEL

$$(5.5) \quad \ln L_i = -\frac{1}{2} \left(\frac{\log r_{e,i} - \beta_{SM} \log M_{\star,i} - \alpha_{SM}}{\sigma_i} \right)^2 - \frac{1}{2} \ln(2\pi\sigma_i^2) - \frac{1}{2} \ln \left(\frac{\log M_{\star,i} - \mu_{SM}}{\sigma_{M,i}} \right)^2 - \frac{1}{2} \ln(2\pi\sigma_{\log M,i}^2) - \ln F_i(r_{e,i})$$

with dispersion for the i^{th} EEL $\sigma_i^2 = \sigma_{SM}^2 + \Delta(\log r_{e,i})^2$ for observational uncertainty $\Delta(\log r_{e,i})$; dispersion of the underlying Gaussian distribution of stellar mass $\sigma_{M,i}^2 = \tau_{SM}^2 + \Delta(\log M_{\star,i})^2$; $F_i(r_{e,i})$ is the relative magnification (i.e. the lensing probability, the dashed line in Figure 5.3) for the i^{th} EEL at radius $r_{e,i}$, and $M_{\star,i}$ and $r_{e,i}$ are measured in units of $10^{11} M_\odot$ and kpc, as before. The first term here is the usual χ^2 term and the second is its normalisation which must be included in the likelihood calculation as it depends on the intrinsic scatter σ_i^2 , which is a model parameter. The third and fourth terms describe the normal distribution of the underlying parent distribution of stellar masses, and the last term accounts for the bias due to lensing.

The lower panel of Figure 5.3 shows our inference on the intrinsic size-mass relation (using the 2C models, but the 1C models yield a consistent result), and the posteriors are summarised in Table 5.5. We find that the intrinsic slope is marginally shallower than the observed slope, and consistent with the $z = 0.75$ van der Wel et al. (2014) slope, and still offset to smaller sizes. It therefore seems that this population of compact galaxies has a size-mass relation which is systematically offset from that of the global population.

5.5.4 Morphologies

As suggested in Section 5.5.2, the massive, compact nature of the EELs sources, together with their intermediate redshifts, may indicate that they are relic red nuggets, or red

nuggets caught in the act of evolving. Either way, the resolving power of lensing allows us to characterise their morphologies in detail and so attempt to distinguish between different models of red nugget evolution (Fan et al., 2010; Hilz et al., 2013; Ishibashi et al., 2013). To this end, in this section we characterise the morphologies of our EELs sources and compare them with those of the global SDSS galaxy population. Following this, we compare them with other red nuggets (the subject of Section 5.6.1) and predictions for red nugget growth (Section 5.6.2).

First, we compare the EELs sources with the global SDSS galaxy population using the bulge+disk decomposition catalogue of Simard et al. (2011). This provides fits to a sample of roughly 1.1 million galaxies from SDSS DR7 using three different models: a pure Sérsic model (equivalent to our 1C models), an $n_{bulge} = 4$ and exponential disk model, and an $n_{bulge} = \text{free}$ and exponential disk model (comparable, but not equivalent, to our 2C models). Specifically, we ask the question, *Do the EELs sources have any distinguishing features relative to the global galaxy population?*

We find that the distributions of axis ratios and Sérsic indices for our 1C models are both consistent with the global population. Though our sample size is small, Kolmogorov-Smirnov (KS) tests in both cases do not allow us to reject the null hypothesis that both the EELs sources and the Simard et al. (2011) galaxies are drawn from the same distribution. We note, however, that all of our sources have $n_{1C} \gtrsim 4$, which seems to indicate that all have significant bulge components – that is, none are purely disk. These two null results are interesting in light of the finding of a high incidence of flattened and disk objects in the Stockton et al. (2014); Hsu et al. (2014) samples of low-redshift red nugget relics, and will be revisited in Section 5.6.1.

On the other hand, we find a much higher proportion of EELs sources needing two-component models relative to that in the Simard et al. (2011) catalogue. First, we note that nine out of thirteen ($\sim 70\%$) of our sources require two-component models in order for the data to be described down to the noise; in contrast, the Simard et al. (2011) catalogue provides a probability $p(Ps)$ that a bulge+disk decomposition is *not* needed over the pure Sérsic model, and indicates that objects with $p(Ps) < 0.3$ may be treated as requiring a bulge+disk decomposition whilst those with $p(Ps) > 0.3$ may be considered spheroidal. We use this to classify the galaxies in their sample and find that only $\sim 20\%$ fall into the bulge+disk category. This is particularly striking given that the Simard et al. (2011) catalogue contains spiral galaxies in addition to ETGs (they do not apply morphological cuts), whereas our EELs sources are all ETGs (though we note that our criteria for adding a second Sérsic component are not the same as Simard et al., 2011’s $p(Ps)$ probabilities). This seems to indicate a significant morphological difference between the ETGs in the

two samples, with our galaxies being much more likely to have a flatter, more extended component in addition to the central bulge. This is further underlined by the distribution of Sérsic indices that we infer for the more extended components of our 2C models, for which $n_{env} < 1$ in all but three cases and $n_{env} < 1.5$ in all but one case. We do not require the more extended Sérsic component to have a lower Sérsic index and it is entirely possible for objects to require two relatively spheroidal components, e.g. oriented at different angles or with particular combinations of n and R_e to reproduce their structures, so the fact that all our 2C models yield a flat component is further evidence that these objects tend to have disks or envelopes surrounding their central cores. This is a finding we will return to in Section 5.6.2 in the context of red nugget growth.

5.6 Discussion

5.6.1 Are the EELs sources red nugget relics?

Previous studies of red nugget morphologies have been carried out at high redshift by van Dokkum et al. (2008), and of intermediate-redshift red nugget relics by Auger et al. (2011), Stockton et al. (2010, 2014) and Hsu et al. (2014). One general finding of the lower-redshift work was that large proportions of their samples required two-component Sérsic models to describe the data satisfactorily, and that these two-component models generally implied disk morphologies. In this respect, our results are in accord: we also find nine out of the thirteen EELs to require two-component models. However, many of the two-component models of Stockton et al. (2010, 2014) and Hsu et al. (2014) differ strikingly from ours in that, for nearly all their objects, the Sérsic indices of *both* components are consistently low – with, for instance, five out of the seven systems in Stockton et al. (2014) having both components with $n < 1.6$. While the Hsu et al. (2014) sample finds more of a range of morphologies – possibly due to their larger sample size – they also classify twelve out of their twenty-two (55 %) objects as disk-like, with only two of these twelve exhibiting convincing bulges. The Sérsic models of van Dokkum et al. (2008) are also in line with this, with their nine objects having generally low Sérsic indices ranging between $0.5 < n < 4.5$ (though they do point out the uncertainty inherent in measuring galaxy structure at those redshifts). This is extremely different from what we report in Section 5.5.4, and suggests a paradigm in which ETGs are originally disk-like and become more spheroidal over time; though we have a large number of galaxies with some kind of outer envelope or disk, these are all accompanied by bulge-like components with $n > 4$, lending themselves very naturally to the interpretation that originally spheroidal galaxies, assembled at high

redshift, have grown by accreting matter at large radii.

This difference is surprising, and may suggest that the EELs sources are not drawn from the same population as these other objects. As those studies were particularly focussed on high-redshift nuggets, or relics of high-redshift nuggets, it may be the case that the EELs sources represent the more evolved counterparts of theirs. On the other hand, the difference may be the effect of our different search methods and selection criteria. In particular, the intermediate-redshift studies used IR photometry in addition to SDSS data in order to identify compact candidates, whereas we extract our compact galaxies from SDSS using strong lensing. Alternatively, it may be the case that the models in these previous studies were subject to larger uncertainties in their structural parameters than thought or that they were systematically underestimated. Indeed, van Dokkum et al. (2008) do note the difficulty in determining the morphologies of such small, distant objects; this is a problem that is still present to some extent at the redshifts probed by Stockton et al. (2014) and Hsu et al. (2014), but which is mostly mitigated in our analysis due to the fact that our sources have been lensed. At this stage, it is not possible to discriminate between these possibilities and so the picture remains complex. What is clear, however, is that compact ETGs at intermediate redshifts have a range of morphologies and may be at different stages in their evolution.

5.6.2 Are the EELs sources evolving red nuggets?

It is possible that the EELs sources are not relic red nuggets, but the descendants of red nuggets, caught in the middle of their evolution. If so, we should be able to interpret their characteristics in the context of red nugget evolution.

We have shown the EELs sources to have Sérsic indices that are generally consistent with the global distribution – though possibly under-representing the low- n tail – when modelled using single components, while two-component models almost always have a low- n component in addition to a bulge. This is at least qualitatively consistent with the simulations of Hilz et al. (2013), which considered the growth of ellipticals via minor mergers and found this to lead to inside-out growth, with the central density remaining relatively unaffected while matter is accreted in the outer parts, such that the bulge becomes embedded in an envelope of accreted matter.

The minor-merger-driven expansion scenario of Hilz et al. (2013) also predicts that the stars added at large radii should be metal-poor. In Section 5.6.3, we find negative colour gradients for nearly all the EELs sources, with the outskirts being bluer than the central regions; however, without spectral information we cannot say whether these gradients are

being driven by age (with younger stars at larger radii) or metallicity (with metal-poorer stars at larger radii). It is therefore difficult to interpret this finding in the context of the action of mergers. Interestingly, one other prediction of those simulations is that the central dark matter fraction should undergo strong evolution with redshift (from $\sim 40\%$ at $z = 2$ to $\gtrsim 70\%$ today); estimating the dark matter fractions of our EELs sources from the stellar kinematics would be a useful further test of this scenario, and is something we consider in Chapter 6.

We note that there are a number of alternative explanations for red nugget growth, including the AGN-feedback-driven scenario proposed by Ishibashi et al. (2013), which allows radiation pressure to trigger star formation at large radii, and the quasar-driven ‘puffing-up’ scenario proposed in Fan et al. (2008), which has the expulsion of gas from the inner regions to the outskirts responsible for size evolution in these systems. These models may also lead to the bulge+envelope morphologies that seem to characterise the EELs sources; however, they do not as of yet make any quantitative predictions that would allow a more direct comparison with our data and we therefore do not comment on them any further here. We emphasise that even our small sample reveals a diversity of morphologies. This may indicate that we are seeing objects at various stages in their evolution, but may also be evidence for the range of evolutionary mechanisms that are at work.

5.6.3 Colour gradients and inside-out growth

A general prediction of the hierarchical formation scenario for massive galaxies is that the gradual accretion of younger, lower-metallicity stars from lower-mass satellites should lead to negative colour gradients across the galaxy, with the central parts generally containing an old but more metal-rich stellar population compared the outskirts. This has been observed in a number of low-redshift ETGs (e.g. Franx et al., 1989; Peletier et al., 1990; Tamura & Ohta, 2003; Kuntschner et al., 2010; Tortora et al., 2010) and also in simulations (De Lucia et al., 2006; Tortora et al., 2013). If red nuggets grow significantly in size, they should represent extreme examples of inside-out evolution. The EELs sources, which may be the descendants of these systems, therefore present an ideal opportunity to test these expectations.

Recently, Tortora et al. (2016) placed the first constraints on colour gradients in compact ETGs using their catalogue of 92 systems at redshifts $z \sim 0.2 - 0.7$, enabled by their high signal-to-noise KiDS dataset, and found preliminary evidence for negative gradients, consistent with the general ETG population. Here, we can exploit the magnification of our nuggets due to lensing to further constrain the colour gradients in our sample.

To do this, we create a new set of lens models in which all three bands are modelled simultaneously. To limit the dimensionality of the inference, we fix the mass profile of the lensing galaxy using our previous models (see Section 5.3), and infer the light profiles of both lens and source and the K' -band PSF. We assume the lens galaxy’s light profile to be the same in each band but we allow the source to have a wavelength-dependent half-light radius described by Equation 5.2. The location, ellipticity and position angle of the source are required to be the same in all bands (though we allow for an offset between bands due to imperfect image registration), and we use a single Sérsic component to allow a straight-forward interpretation of the wavelength dependence of the radius.

We find that ten out of twelve of the sources that we were able to successfully model exhibit clear negative gradients, with a sample median $\alpha = -0.45$ (and standard deviation $\sigma = 0.08$); of the remaining two objects, one (J1347) has a gradient consistent with zero and the other (J1144) has a mildly positive gradient. We were not able to find a satisfactory model for J1619 (see Appendix D) and exclude it from the analysis. A range of gradients – mostly negative, but some positive – was also noted by Tortora et al. (2016) and taken to indicate the range of initial conditions which can enable such objects to form; the properties of our sample underline this result, though we suggest that it may also indicate the diversity of evolutionary paths that these systems can follow.

Interestingly, one object (J1125) has an extremely large negative gradient $\alpha_R = -1.83 \pm 0.11$, indicative of extreme changes in the stellar population as a function of radius, and therefore, potentially, a very extended period of accretion. We note that our 1C model for J1125 had a high Sérsic index in the K' band ($n = 8.40 \pm 0.98$) as compared to the HST bands ($n = 6.24 \pm 0.23$), which is consistent with a picture of the bulge being especially bright in the red, with faint, extended wings, and less bright at blue wavelengths relative to the wings. The very compact bulge size in J1125’s 2C model is also interesting, and it may be that we are seeing an extreme case of inside-out growth in this system.

Finally, we investigate the correlations of the colour gradients – characterised by α_R – with redshift z_s and stellar mass M_* , in each case modelling the correlation as $\alpha_R = \alpha_R X + b_R$ for variable X (i.e., the redshift or stellar mass) being drawn from a Gaussian distribution with mean μ_X and standard deviation τ_X . As shown in Figure 5.4 (left), we find the colour gradients to be weaker (α less negative) at higher redshifts, and suggest that this may be because colour gradients become imprinted over time as more inside-out growth takes place. We also find that the colour gradients are weaker at higher stellar masses (right panel); this may be the result of stellar populations in merger events mixing being more efficient at higher masses (Kobayashi, 2004, but also see Tortora et al., 2009 for a suggestion that strong quasar feedback at high redshifts could be responsible

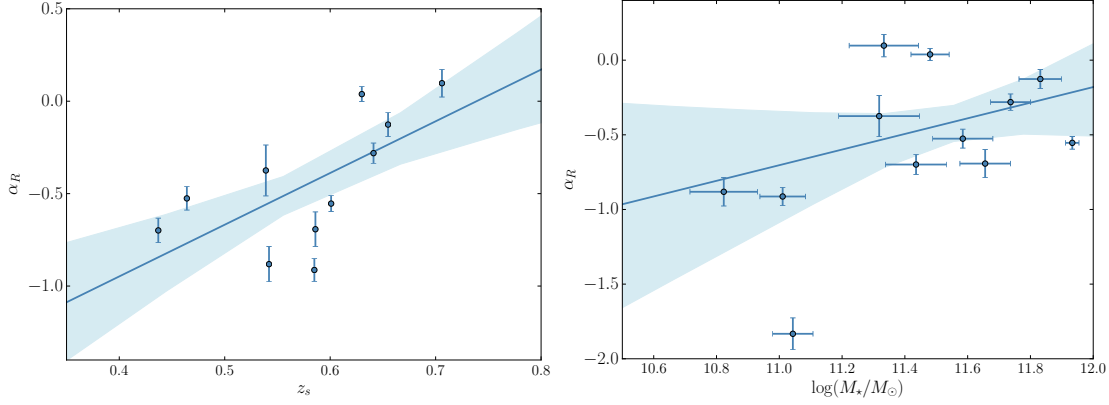


Figure 5.4: Inference on relations between the radial colour gradients α_R and source redshift (left) and stellar mass (right). In both cases, we model the data as falling on a linear relation $\alpha_R = aX + b$ with some intrinsic scatter, with variable X drawn from a Gaussian distribution with mean μ and standard deviation τ . In both cases, we find a general trend that the radial gradients become stronger at low redshift and low mass.

for flattening out the colour gradients in high-mass galaxies).

5.6.4 Growth in dense environments?

A number of low-to-intermediate-redshift studies of red nuggets have suggested an important role for environment in the formation and survival of massive compact galaxies. Stringer et al. (2015) used cosmological simulations to track the evolution of a sample of compact systems and found that 94% became associated with larger structures – either ending up embedded in clusters, or passing through such structures at an earlier phase in their lifetimes. From an observational point of view, Valentinuzzi et al. (2010a,b) identified a significant fraction of cluster galaxies as massive and compact, both at $0.4 < z < 1.0$ and locally at $0.04 < z < 0.07$, while Poggianti et al. (2013) found the fraction of nuggets in the field at $0.03 < z < 0.11$ to be a factor of three smaller than this (though we note that the compactness criterion used in Poggianti et al. (2013) is stricter than that in Valentinuzzi et al. (2010a,b)). On the other hand, more general studies of ETG size evolution (i.e. not limited to compact systems) remain conflicted as to the importance of environment for the ETG population as a whole (e.g. Huertas-Company et al., 2013; Lani et al., 2013; Delaye et al., 2014; Newman et al., 2014). The question arises, then, as to whether we are able to characterise the environments of the EELs source galaxies.

To that end, we investigate the SDSS galaxy population in the regions local to each

source. We note here that the synthetic (i.e., as determined by the stellar population modelling of Section 5.5.1) *gri* magnitudes for all thirteen EELs sources imply that they would have been detected in the SDSS *r* and *i* bands even if they hadn't been lensed. For each object, we query the SDSS photometric database to identify all galaxies with projected separations less than 1.5 Mpc, photometric redshifts within 0.01 of the source redshift and $0.5 < \chi^2 < 2$ for the chi-squared value of the photometric redshift; the last criterion is intended to remove objects with rogue redshifts from our count. For each source, we thus obtain an estimate of the number of galaxies which could be associated with it or become associated with it later on. We then query the database to compile catalogues of objects with similar properties to each EEL, this time using the same redshift criteria but requiring colours in the *gri* bands to be within 0.2 magnitudes of our synthetic SDSS magnitudes for the source and imposing no cut on right ascension and declination. We call these the 'twin' catalogues, and limit each one to 1000 objects. We then repeat the first step for each object in each of the twin catalogues, querying the SDSS database to estimate the number of possibly associated galaxies. This allows us to compare the distribution of associated galaxies for objects in an EEL's twin catalogue with the number of associated galaxies for the actual EEL, and so determine whether or not the EEL is residing in a particularly under- or over-dense environment with respect to other similar galaxies. We find all the EELs to be consistent with their twin catalogues, suggesting that their environments are typical of other SDSS galaxies at similar redshifts and with similar intrinsic SEDs and luminosities. This is in contrast to the suggestions of e.g. Valentinuzzi et al. (2010a) (though see Morishita et al., 2016 for a recent review), though we note that we cannot put strong constraints on this using photometric redshifts alone.

5.7 Summary and conclusions

A great deal of effort has gone into explaining the evolution of compact, massive ETGs at high redshifts into the ETGs that we see in the local Universe. Proposed physical mechanisms for this growth include repeated minor merging and radiative or gas-driven AGN feedback (Fan et al., 2010; Hilz et al., 2013; Ishibashi et al., 2013), each of which makes particular predictions for the way in which these objects should evolve structurally. One of the current challenges is to identify compact objects at intermediate redshifts with which to test these predictions. We have presented a new class of ETG/ETG lenses, the EELs, and have used multiband photometry, exploiting the magnifying effect of lensing, to model the source galaxies with unprecedented resolution. These galaxies form a population of massive, compact galaxies at redshifts $z \sim 0.4 - 0.7$, and may therefore be intermediate-

redshift relics of high-redshift nuggets or partly-evolved nuggets. We have carried out a survey of their structural properties so as to compare with the predictions of various models for red nugget evolution, as well as with other known or candidate low-redshift compact galaxies. Our general findings are as follows.

1. The EELs sources form a massive, compact galaxy population at redshifts $z \sim 0.4\text{--}0.7$, lying systematically below the size-mass relation of ETGs at these redshifts.
2. Generally, two Sérsic components are needed to fully characterise their surface brightness distributions. This indicates complex (though smooth) morphologies and the presence of a bulge-like central component alongside a much lower- n envelope-like component, both of which are compact. Indeed, two out of our thirteen objects have clear, compact envelopes. These may be the result of ongoing accretion onto the compact cores which are already in place at high redshift, in line with an inside-out formation scenario via repeated minor mergers. The diversity of structures that we observe in our small sample highlights the strong evolution that these objects undergo at intermediate redshifts.
3. The EELs sources generally exhibit negative colour gradients, with redder centres and bluer outskirts. While we cannot disentangle the contributions from the age and metallicity of the stellar populations, we note that accretion of lower-mass galaxies with younger or lower-metallicity stars would be consistent with this trend. We also find that colour gradients are stronger at lower redshift and lower stellar mass, in line with a picture in which low-redshift galaxies have experienced more accretion and high-mass galaxies are more efficient at mixing their stellar populations.
4. The EELs sources do not appear to occupy over-dense environments with respect to other SDSS galaxies with similar colours, luminosities and redshifts. This is contrary to suggestions that compact galaxies eventually become embedded in groups or clusters, though we cannot place strong constraints on this at present.

The lensing of these compact galaxies allows us to model their structures in detail and so place constraints on scenarios for their evolution. As low-redshift relics start to be discovered in increasing numbers, these constraints will be valuable in order to understand the evolving number density of these objects and the implications of this on our understanding of the local Universe. Furthermore, additional clues to their evolutionary history will be uncovered with spectroscopic observations to constrain the dynamics and stellar populations of these galaxies, such as those presented in the following chapter.

THE FUNDAMENTAL PLANE OF EVOLVING RED NUGGETS

Abstract

We present an exploration of the mass structure of a sample of 12 strongly lensed massive, compact early-type galaxies at redshifts $z \sim 0.6$ to provide further possible evidence for their inside-out growth. We obtain new ESI/Keck spectroscopy and infer the kinematics of both lens and source galaxies, and combine these with existing photometry to construct (a) the fundamental plane of the source galaxies and (b) physical models for their dark and luminous mass structure. We find their fundamental plane to be tilted towards the virial plane relative to the local fundamental plane, and attribute this to their unusual compactness, which causes their kinematics to be totally dominated by the stellar mass as opposed to dark matter; that their fundamental plane is nevertheless still inconsistent with the virial plane implies that both the stellar and dark structure of ETGs is non-homologous. We also find the intrinsic scatter of their fundamental plane to be comparable to the local value, indicating that variations in the stellar mass structure outweigh variations in the dark halo in the central regions of early-type galaxies. Finally, we show that inference on the dark halo structure – and, in turn, the underlying physics – is sensitive to assumptions about the IMF, but that physically-motivated assumptions about the IMF imply haloes with sub-NFW inner density slopes, and may present further evidence for the inside-out growth of compact early-type galaxies via minor mergers and accretion.

6.1 Introduction

The discovery that massive, passive galaxies at redshifts $z \sim 2$ are much more compact than their present-day counterparts (Daddi et al., 2005; Trujillo et al., 2006; van Dokkum et al., 2008) has led to a picture in which ETGs evolve dramatically in size over the course of their lives. Moreover, the detection of extended outer envelopes surrounding lower-redshift compact ETGs (e.g. van Dokkum et al., 2010) implies an important role for dissipationless merging and accretion in evolving these systems towards the present-day size-mass relation. However, though a consensus is now building over the evolution of their luminous structure, their dark halo structure remains elusive. Even in normal (non-compact) ETGs, how the dark halo is affected by baryonic processes such as mergers and accretion is not well understood; simulations suggest that dynamical heating from infalling satellites should displace dark matter from the centre of the halo to larger radii (El-Zant et al., 2004; Laporte et al., 2012), but this must compete with other processes such as adiabatic contraction during the infall of stellar material and feedback from supernovae and AGN (Read & Gilmore, 2005; Martizzi et al., 2013). Observationally, the picture is also unclear, with the halo structure of ETGs exhibiting a diversity which may depend on environment (Newman et al., 2015). Probing the mass structure of low-redshift, partly-evolving massive compact ‘red nugget’ galaxies, where the evolution of the baryonic material is dominated by merging, allows us to isolate this aspect of inside-out growth from other baryonic processes and investigate its impact on the haloes of individual galaxies much more closely.

The mass structure of ETGs *as a population*, on the other hand, has historically been accessed through the fundamental plane which tightly relates their characteristic size, velocity dispersion and surface brightness (Djorgovski & Davis, 1987; Dressler et al., 1987). The existence of such a plane follows directly from the assumption that galaxies are virialised – and, to some extent, homologous – systems (see e.g. Ciotti et al., 1996), with the small intrinsic scatter indicating a strong degree of regularity in their formation and evolution. However, the fact that the fundamental plane is tilted relative to the virial prediction implies some degree of non-homology, with mass-dependent variations in either the luminous matter – for instance, the IMF, stellar mass structure and stellar dynamics – or the dark matter – including the dark halo structure, concentration and the dark matter fraction – or both; however, it is difficult to disentangle effects of variations in the dark and light structure (see e.g. Trujillo et al., 2004; Cappellari et al., 2006; La Barbera et al., 2008) so as to extract information on galaxy structure (though the construction of the *mass plane* has helped to discount dynamical non-homology as the main cause of the fundamental plane tilt; see Bolton et al., 2008). Constructing the fundamental plane for

compact galaxies, whose luminosity-weighted velocity dispersions probe the very central regions where the dark matter fraction is expected to be low, provides a way to separate the contributions to the fundamental plane tilt from the stars and the halo and so better understand the mass properties of the ETG population.

In this chapter, we present new high signal-to-noise ESI/Keck spectroscopy for the 13 EELs introduced in Chapter 5, and combine these data with photometry to probe the mass structure both of individual galaxies and of the population as a whole. The background galaxies of the EELs are massive and compact ETGs, which the combination of lensing magnification and high-resolution imaging data can resolve on ~ 100 pc scales, making them the ideal sample with which to probe ETG structural evolution. In Section 6.2, we introduce the data, the data reduction and the kinematic modelling. Section 6.3 presents the fundamental plane; in Section 6.4 we construct physical models of the EELs sources in order to set constraints on their dark matter content, and in Section 6.5 we discuss our findings and conclude.

6.2 Data and kinematic modelling

We observed the 13 EELs using the Echelette Spectrograph and Imager (ESI; Sheinis et al., 2002) on Keck on the nights of 2013 May 14, 2015 Jan 23 and 2016 July 08, obtaining 1-hour exposures for each system (except J1218+5648, J1605+3811 and J2228-0018, which were observed for 30 minutes, 10 minutes and 10 minutes respectively), using a slit width of $0.75''$. The data were reduced with a custom-made, python-based pipeline and the wavelength scale calibrated using arc lamp exposures taken on the night. For each system, we extracted spectra over two separate apertures – one centred on the lens, with width $0.5''$, and a second centred on the brightest part of the Einstein ring, with width $0.3''$ – to obtain spectra which maximised the relative signal from the lens and the source respectively. For one system (J1606+2235), the slit did not cover the surface brightness peaks in the source galaxy which meant that we were not able to extract a spectrum of the central regions of the source in this case; we therefore exclude it from our sample. The (source-galaxy-centred) spectra for the 12 remaining systems are presented in Figure 6.1.

To determine the stellar velocity dispersions, we model each spectrum as the sum of a lens, source and additive continuum component. For the lens and source, we use stellar templates for A, F, G and K stars from the Indo-US Stellar Library of Coudé Feed Stellar Spectra (Valdes et al., 2004), which we redshift and convolve with a dispersion $\sigma_{model}^2 = \sigma_{true}^2 + \sigma_{inst}^2 - \sigma_{tmp}^2$ where σ_{true} is the physical velocity dispersion of the system, $\sigma_{inst} = 20.30 \text{ kms}^{-1}$ is the instrument resolution and σ_{tmp} is the intrinsic resolution of the

templates (which is 1.2 \AA for the Indo-US templates). The continuum is an additive order-6 polynomial which accounts for the difference in shape between the templates and the true spectrum, and regions where atmospheric absorption dominates the spectrum are masked. We therefore have four free non-linear parameters – the redshift and velocity dispersion for each of the lens and the source – and 24 linear parameters – the weights of each of the nine stellar templates for the source and the lens (4 K stars, 3 G stars and one A and F star each), and the coefficients of the order-6 polynomial – which we explore using MCMC sampling. Our kinematic models are also shown in Figure 6.1 and the kinematics for the sources, together with their photometric properties, are summarised in Table 6.1. We defer the presentation of the lens galaxy kinematics to Chapter 8.

We test the robustness of our kinematic inference by repeating the modelling process using the lower-resolution galaxy templates of Bruzual & Charlot (2003) and find that the typical uncertainty in the velocity dispersion of the source is of order 5%; as this is larger than our statistical uncertainties, we impose this as the uncertainty on all our velocity dispersion measurements. We also check the robustness of our method by modelling the kinematics of (a) the ESI spectra, extracted over apertures of width $1.5''$, centred on the lens (to emulate the SDSS apertures, which are circular with radius $1.5''$) and (b) the actual SDSS spectra, and compare our inference on the lens kinematics in each case; the uncertainty indicated by these tests is typically smaller than the uncertainty due to the choice of templates.

We combine these kinematics with the photometry presented in Chapter 5 in order to construct both the fundamental plane and physical mass models. In that study, parametric light and lensing mass distributions were used to model the HST/ACS V/I and Keck/NIRC2 K' imaging of the thirteen EEL systems. We refer the reader to that chapter for full details on the modelling and results, but note that here we use the photometric models which allow the source galaxies to have two Sérsic components (labelled ‘2C models’ in that chapter) as, for some systems, the one-component models do not provide an adequate description of the light profile (leading to large differences between one-component and two-component size measurements for a small number of objects). We summarise the effective radii R_e and effective surface brightnesses $\log I_e$ (defined as the average surface brightness within the effective radius, evaluated in the rest-frame Johnson V band using the photometric models of Chapter 5) in Table 6.1.

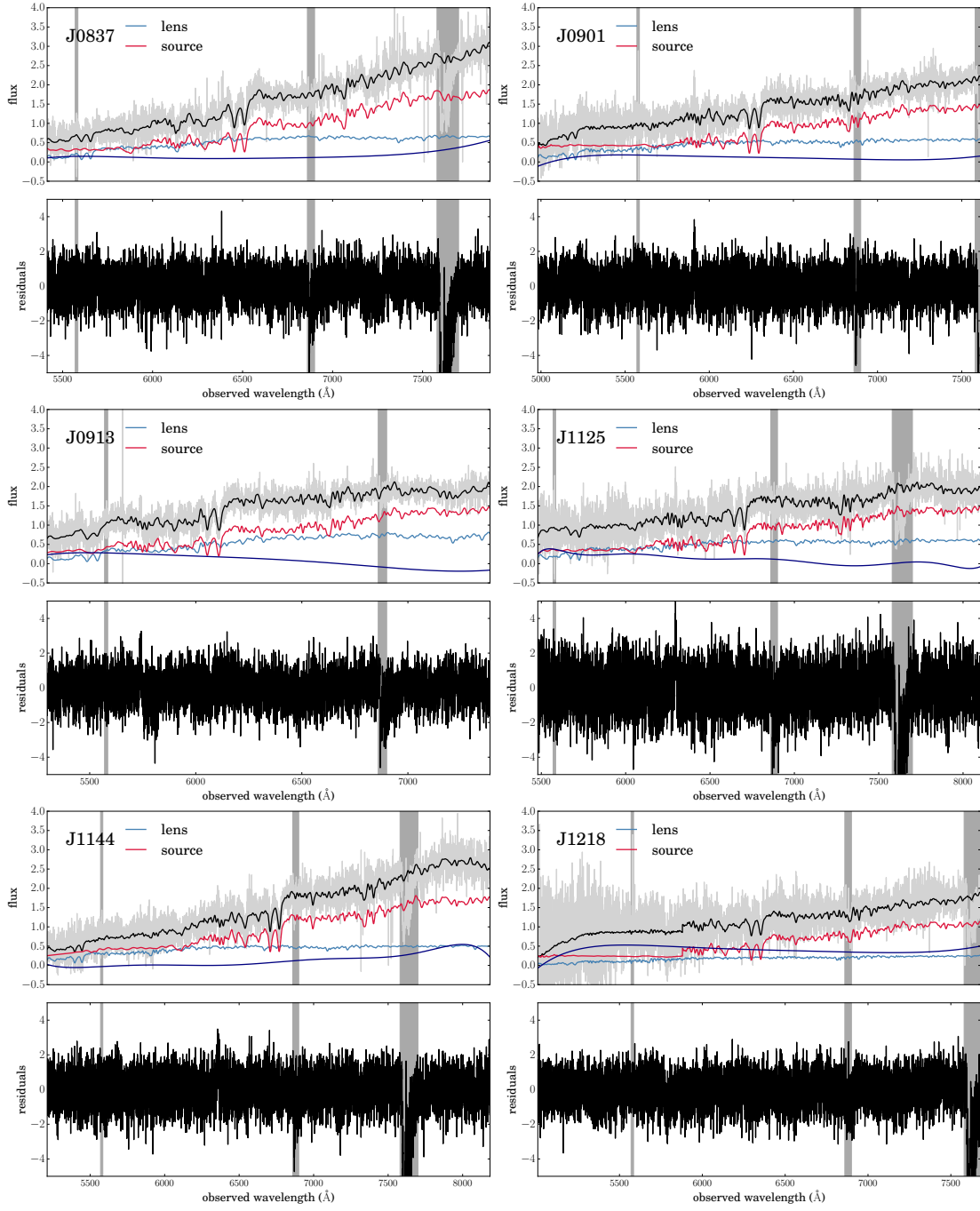


Figure 6.1: Keck/ESI spectra for the 12 EELs in our sample, with kinematic models overplotted. In the upper panels, the data are shown in grey, the total model spectrum in black, and the contributions to the model from lens, source and continuum in blue, red and purple respectively. The lower panels show the residuals of the kinematic models; vertical grey bands represent regions that were masked from the fit due to the presence of telluric absorption features.

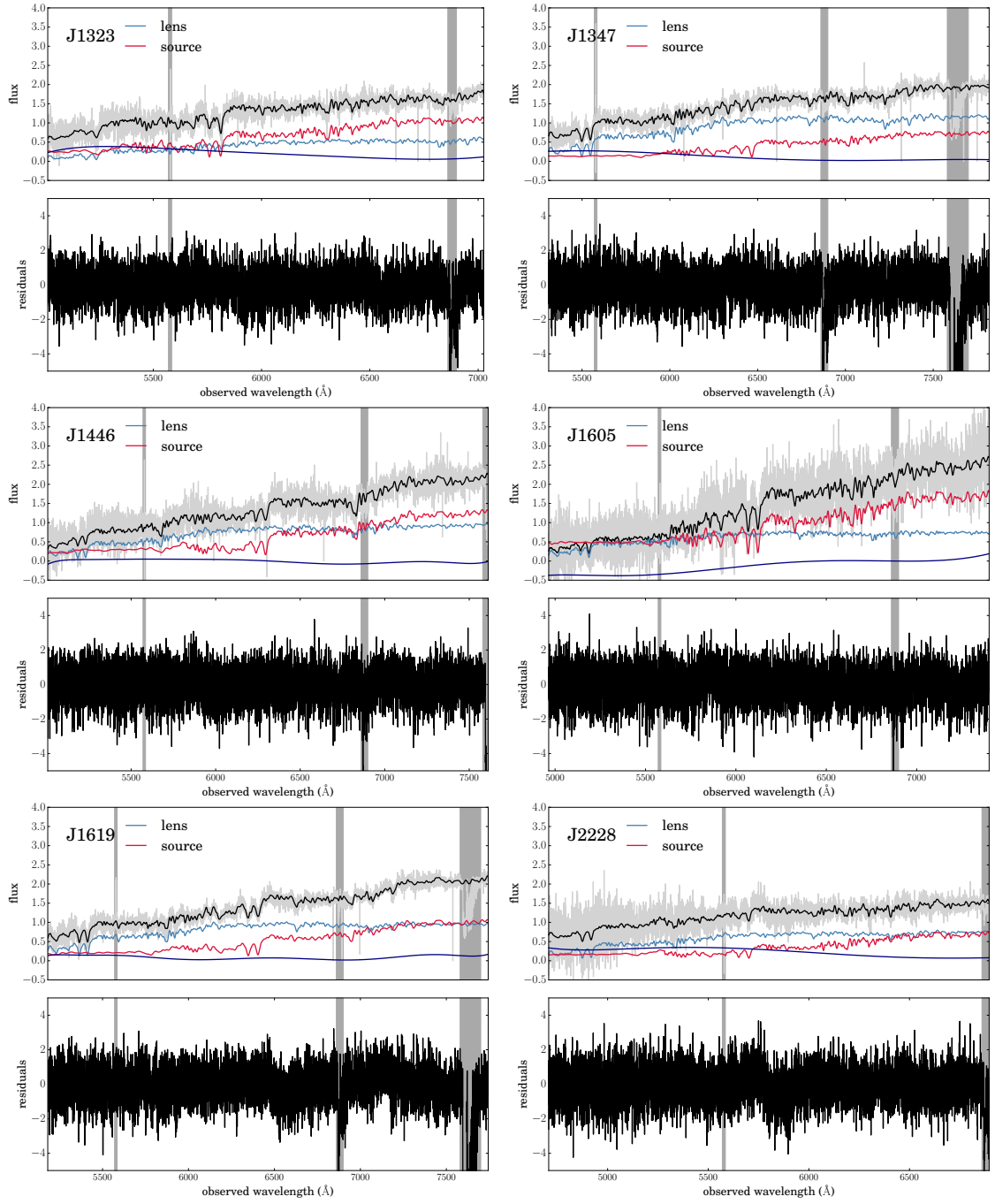


Figure 6.1: continued.

EEL	R_e (kpc)	σ (kms $^{-1}$)	$\log I_e$ (L $_{\odot}$ kpc $^{-2}$)
J0837+0801	4.42 ± 0.27	253 ± 13	9.14 ± 0.01
J0901+2027	3.26 ± 0.19	205 ± 10	8.89 ± 0.01
J0913+4237	4.11 ± 0.17	193 ± 10	8.86 ± 0.01
J1125+3058	1.17 ± 0.02	182 ± 9	9.61 ± 0.01
J1144+1540	9.64 ± 0.28	225 ± 11	8.50 ± 0.02
J1218+5648	6.79 ± 0.33	191 ± 10	8.67 ± 0.01
J1323+3946	1.82 ± 0.11	162 ± 8	9.14 ± 0.01
J1347-0101	5.39 ± 0.49	152 ± 8	8.33 ± 0.02
J1446+3856	1.59 ± 0.04	207 ± 10	9.49 ± 0.01
J1605+3811	2.56 ± 0.05	160 ± 8	9.08 ± 0.01
J1619+2024	5.24 ± 0.20	283 ± 28	8.95 ± 0.01
J2228+2024	4.15 ± 0.08	149 ± 15	8.81 ± 0.01

Table 6.1: fundamental plane data for the EELs source galaxies. Photometry is evaluated in the rest-frame Johnson V band using the photometric models presented in Chapter 5, and kinematics are measured as described in Section 6.2.

6.3 Fundamental plane

6.3.1 The observed fundamental plane

We model the fundamental plane for the twelve EELs sources as

$$(6.1) \quad \log R_e = \alpha \log \sigma + \beta \log I_e + \gamma$$

for effective radius R_e in kpc, velocity dispersion σ in units of 200 kms $^{-1}$ and effective surface brightness I_e in 10^9 L $_{\odot}$ /kpc 2 .

We follow the formalism presented by Kelly (2007) by modelling both $\log \sigma$ and $\log I_e$ as being drawn from multivariate Gaussian distributions with some mean μ and dispersion $\tilde{\tau}$, and allowing some intrinsic scatter σ_{FP} along the direction of $\log R_e$; regarding the ongoing discussion on how to fit the fundamental plane (Hyde & Bernardi, 2009), we note that this method is closer to the case of minimising the residuals along $\log R_e$ rather than perpendicular to the plane. Full details on the form of the likelihood function can be found in Kelly (2007), but essentially, we construct the likelihood for the data given a particular set of plane and parent distribution parameters, and obtain the posterior distribution using an MCMC exploration of the parameter space. The fundamental plane that we infer is shown in both side-on and face-on projections in Figure 6.2, and our inference on the parameters is summarised in Table 6.2. As our sample is small, we test the robustness of our inference by remodelling random subsamples of the dataset and find that the uncertainty introduced by this is small, and less than half the size of the statistical

uncertainties. We find that the observed fundamental plane of the EELs sources is steep relative to the fundamental plane of ‘normal’ galaxies, with $\alpha = 1.54 \pm 0.36$ as compared to the $\alpha = 1.24 \pm 0.07$ that is found locally, and $\beta = -0.68 \pm 0.06$ compared to the local value $\beta = -0.82 \pm 0.02$ (Jørgensen et al., 1995); note though that both are consistent with the ‘normal’ fundamental plane parameters within 2σ . As a reference we use the sample of 17 ETGs belonging to the $z = 0.545$ cluster MACSJ0717.5+3745 which were presented in Chapter 4; these ETGs have a similar redshift and mean velocity dispersion to those in the EELs sample, but fundamental plane parameters that are consistent with the local values. Figure 6.2 shows that these do indeed fall differently on the fundamental plane. We also note that Auger et al. (2010a) found that the inclusion of an intrinsic scatter around the fundamental plane further tilts the plane of normal ETGs along the $\log \sigma$ such that it is shallower than the Jørgensen et al. (1995) value (Auger et al., 2010a find $\alpha = 1.02 \pm 0.20$, 1.19 ± 0.14 with and without intrinsic scatter respectively). We also find this to be the case, with the removal of an intrinsic scatter from our model giving $\alpha = 1.76 \pm 0.29$ (and having negligible effect on β).

The fact that these ETGs have compact light profiles means that their luminosity-weighted velocity dispersions probe only their very central regions, where the dark matter fraction is minimal. To zeroth order, we might then expect that any deviation in their fundamental plane from the virial prediction must be due to variations in the stellar mass and dynamical structure across the plane, as opposed to variations in the dark matter. That we find their fundamental plane to be tilted relative to both the local ETG population *and* the virial plane indicates that the properties of both the stellar mass *and* the dark matter vary across the plane. However, before drawing conclusions from this, we must account for the effect of the selection function of the EELs sources on the fundamental plane parameters.

	α	β	γ	μ_σ	μ_{I_e}	τ_σ	τ_{I_e}	$\log \sigma_{FP}$
observed	1.54 ± 0.36	-0.68 ± 0.06	0.53 ± 0.02	0.00 ± 0.03	-0.03 ± 0.14	0.07 ± 0.02	0.38 ± 0.11	-1.31 ± 0.19
	1.76 ± 0.29	-0.69 ± 0.04	0.56 ± 0.01	-0.02 ± 0.02	-0.04 ± 0.13	0.06 ± 0.02	0.39 ± 0.11	–
intrinsic	1.41 ± 0.35	-0.66 ± 0.06	0.56 ± 0.02	-0.02 ± 0.03	-0.02 ± 0.12	0.08 ± 0.02	0.35 ± 0.10	-1.28 ± 0.18
	1.69 ± 0.28	-0.69 ± 0.04	0.56 ± 0.01	-0.01 ± 0.02	-0.03 ± 0.12	0.07 ± 0.02	0.37 ± 0.11	–

Table 6.2: Parameters and associated uncertainties for the fundamental plane, modelled according to the equation $\log R_e = \alpha \log \sigma + \beta \log I_e + \gamma$, with intrinsic scatter σ_{FP} about $\log R_e$. The independent variables $\log \sigma$ and $\log I_e$ are modelled as being drawn from Gaussian distributions with mean $\vec{\mu} = (\mu_\sigma, \mu_{I_e})$ and variance $\vec{\tau}^2 = ((\tau_\sigma^2, \rho \tau_\sigma \tau_{I_e}), (\rho \tau_\sigma \tau_{I_e}, \tau_{I_e}^2))$. Velocity dispersions, surface brightnesses and effective radii are measured in units 200 km s^{-1} , $10^9 L_\odot$ and kpc respectively. The first and second rows represent models with and without an intrinsic scatter (the ‘observed’ fundamental plane); allowing for an intrinsic scatter makes the fundamental plane shallower in the $\log \sigma$ directions; the third and fourth rows represent the same, but after accounting for the bias introduced by the EELs source selection function (the ‘intrinsic’ fundamental plane). The selection has a negligible effect on the fundamental plane parameters.

6.3.2 The intrinsic fundamental plane

The EELs sources form a biased population, as they were found in a lens search, and this may have a non-trivial effect on the orientation of the fundamental plane. We now follow the reasoning presented in Chapter 5 to correct for this bias and so recover the *intrinsic* fundamental plane of these compact systems.

The selection function of the EELs sources has three main contributions. Firstly, the source must be lensed by the foreground object; this relates to the cross-section for lensing. Secondly, the inclusion of an EEL in the SDSS spectroscopic sample requires the lens+source system as a whole to satisfy the criteria of the SDSS target selection (Strauss et al., 2002), which itself is non-trivial, though the main effect here is that the system is bright. Finally, the EEL must pass our spectroscopic search, which is subjective but imposes criteria such as the (magnified) source galaxy flux being comparable to the lens galaxy flux and the redshifts of the two objects approximately satisfying $0.1 \lesssim z \lesssim 0.7$. The combination of these conditions leads to some selection function that modifies the intrinsic population of compact galaxies to the population of EELs sources that we observe.

Of these three contributions, the latter two are difficult to quantify and should not introduce any large bias into our measurement of the fundamental plane, although they will push us to the high-surface-brightness end. On the other hand, the first – the lensing cross section – introduces a selection function such that we are relatively more efficient at selecting compact galaxies at lower velocity dispersions. We can understand this as follows: differential magnification introduces a bias towards smaller objects (closer to the line-of-sight of the lens), whereas, for a given size, there is no bias as a function of luminosity, and therefore velocity dispersion (assuming the latter is dominated by stellar mass). The result of this is that an object of fixed velocity dispersion becomes increasingly likely to be seen in the lensed population relative to the intrinsic population as it becomes more compact.

This bias enters the fundamental plane in two ways, making its orientation biased towards objects with (a) small R_e and (b) large $\log I_e$. Following Chapter 5, we modify our likelihood for the fundamental plane data for the i^{th} EEL, given a model, by a function describing the magnification achieved by the lens as a function of source size $F_i(R_{e,i})$, which serves as a proxy for the probability of the source being identified in the lens search. The result is summarised in Table 6.2 for models with and without intrinsic scatter, and shows that the bias is negligible – probably because of the low intrinsic scatter, which makes selection effects unimportant – and the fundamental plane of these compact systems remains tilted relative to that of normal ETGs. As mentioned in Section 6.3.1, this seems to

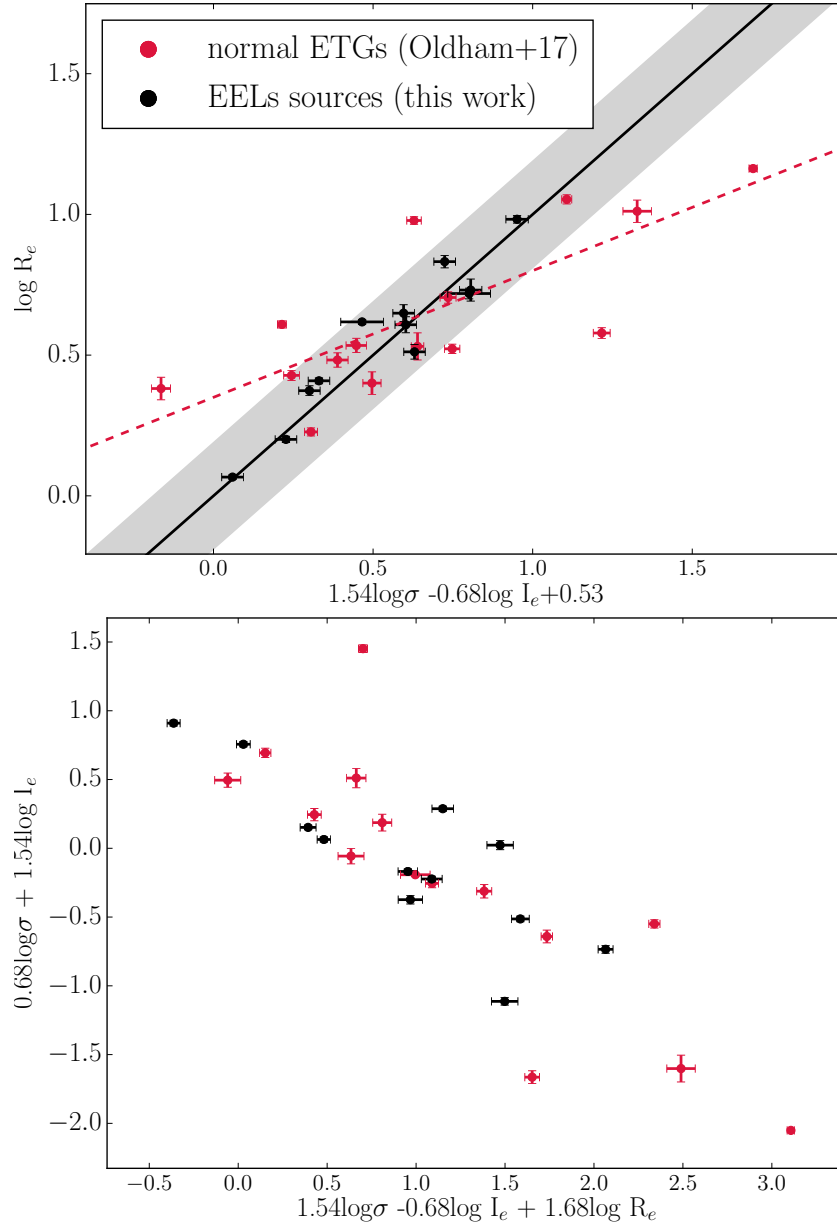


Figure 6.2: The observed fundamental plane of the massive compact EELs sources, shown in side-on (top) and face-on (bottom) projections, with the similar-redshift, less compact ETGs from Chapter 4 shown in red for comparison. The fundamental plane is steeper in $\log\sigma$ and shallower in $\log I_e$ relative to that of the ‘normal’ galaxies, which indicates that the both the dark and stellar mass structure and stellar populations vary as a function of the plane parameters. The fundamental plane also has a comparable intrinsic scatter to that of normal ETGs, implying that the intrinsic scatter is driven by stellar population properties. Velocity dispersions, surface brightnesses and effective radii are measured in units 200kms^{-1} , $10^9 L_\odot$ and kpc respectively.

indicate that the structure of both the dark and luminous mass components varies across the fundamental plane, i.e. as a function of mass. A discussion of this result is presented in Section 6.5.1.

6.4 Physical models

We can also combine our kinematic measurements, which give a measure of the *total* mass in the central regions, with our photometric measurements from Chapter 5, which give a measure of the *stellar* mass, to make inference on the central dark matter fraction. By further supplementing these measurements with abundance matching relations, we can develop toy models to reconstruct the dark halo on large scales and so investigate the dark structure. We emphasise that our data are only sensitive to the mass in the *central* regions, and that strong assumptions about how the stellar mass (which we can measure) relates to the virial mass (which we cannot measure) must be made. Nevertheless, abundance matching relations have been shown to be robust (Behroozi et al., 2010), and can therefore provide useful insight here, allowing us to probe the halo structure out to relatively high redshifts.

Our physical models are constructed as follows. For each EEL, we take the model of the lens system from Chapter 5 and use the surface brightness profile $I(R)$ for the source to construct its stellar mass profile, assuming a stellar mass-to-light ratio which is set by the total magnitude and total stellar mass (under the assumption of either a Chabrier or a Salpeter IMF). We use a generalised Navarro-Frenk-White (gNFW) profile to construct the dark matter halo as

$$(6.2) \quad \rho(r) = \frac{\rho_0}{\left(\frac{r}{r_s}\right)^\xi \left(1 + \frac{r}{r_s}\right)^{3-\xi}}$$

where the normalisation $\rho_0(M_\star)$ is drawn from a normal distribution based on the $z = 0.6$ $M_{halo}(M_\star)$ abundance matching relations of Behroozi et al. (2013, and M_\star is the stellar mass assuming a Chabrier IMF, which was also used in the construction of these abundance matching relations); the inner density slope ξ is drawn from a normal distribution $N(\mu_\xi, \tau_\xi^2)$ such that μ_ξ characterises the mean inner density slope of the EELs source population and τ_ξ represents their scatter and the scale radius r_s is either calculated from the mass-concentration relation (Macciò et al., 2008, though we note that this relation was constructed assuming NFW haloes, and it is not clear that it should still be valid when the NFW assumption is relaxed) or calculated from concentration distribution $N(\mu_c, \tau_c^2)$ with μ_c, τ_c to be inferred. Though the NFW profile is commonly used due to its success

in describing the dark haloes of galaxies in dark-matter-only simulations (Navarro et al., 1997), the use of the gNFW is motivated by the expectation that baryonic physics should modify the halo in some way – either contracting the halo via adiabatic processes, creating a super-NFW cusp $\xi > 1$ (Blumenthal et al., 1986; Gnedin et al., 2004), or hollowing it out via heating from active galactic nuclei or dynamical friction, leading to a sub-NFW central region $\xi < 1$ (e.g. El-Zant et al., 2004; Mashchenko et al., 2006; Governato et al., 2012; Laporte et al., 2012; Velliscig et al., 2014).

We then use the total density profile, $\rho(r) = \rho_{DM}(r) + \rho_*(r)$, to calculate a velocity dispersion using the spherical Jeans equation

$$(6.3) \quad \frac{d}{dr}(l\sigma_r^2) + 2\frac{\beta(r)}{r}l\sigma_r^2 = l(r)\frac{GM(r)}{r^2}$$

where $l(r)$ is the luminosity density of the stars (i.e. the deprojected surface brightness profile, assuming axisymmetry), $\beta(r) = 1 - \sigma_t^2/\sigma_r^2$ is the anisotropy parameter and $\sigma_r(r)$ the radial velocity dispersion, which we then project along the line of sight as

$$(6.4) \quad \frac{1}{2}I(R)\sigma_{los}(R)^2 = \int_R^\infty \frac{l\sigma_r^2 r dr}{\sqrt{r^2 - R^2}} - R^2 \int_R^\infty \frac{\beta l\sigma_r^2 dr}{r\sqrt{r^2 - R^2}}$$

to obtain the line-of-sight velocity dispersion $\sigma_{los}(R)$ as a function of projected radius R . Finally, we integrate σ_{los} over a circular aperture of radius $1.5R_e$ (comparable to the effective aperture over which the lensed source spectra were extracted) to obtain the velocity dispersion that would be measured within the aperture

$$(6.5) \quad \sigma_{ap}^2 = \frac{\int_0^{R_{ap}} I\sigma_{los}^2 R dR}{\int_0^{R_{ap}} I R dR}$$

(see e.g. Mamon & Łokas, 2005). We investigate both isotropic and constant-anisotropy models across the range $-2 \leq \beta \leq 1$, and find that the difference is sufficiently small that our data cannot distinguish between them; we therefore adopt $\beta = 0$ in all models. Thus, given the luminous structure of an EELs source, we can reconstruct the halo and compare the implied central (dark+light) mass with the central velocity dispersion to constrain the dark matter structure in the inner regions.

Our inference is shown in Figure 6.3 and Table 6.3. We find that our inference on the halo inner slope depends strongly on the IMF that is assumed, with bottom-heavy (Salpeter) and bottom-light (Chabrier) IMFs requiring haloes that are excavated and contracted respectively. This is a common problem in attempts to disentangle the dark and luminous mass structure of ETGs and in general must be broken through the use of multiple tracer populations or other mass probes. In this case, the construction of stellar

IMF	free	μ_ξ	τ_ξ	μ_c	τ_c
Chabrier	ξ	$1.23^{+0.08}_{-0.09}$	$0.29^{+0.05}_{-0.08}$	—	—
	ξ, c	$1.15^{+0.11}_{-0.12}$	$0.32^{+0.11}_{-0.06}$	$5.34^{+0.86}_{-1.10}$	$0.51^{+0.45}_{-0.34}$
Salpeter	ξ	$0.71^{+0.13}_{-0.13}$	$0.45^{+0.12}_{-0.08}$	—	—
	ξ, c	$0.65^{+0.14}_{-0.11}$	$-0.46^{+0.09}_{-0.07}$	$4.10^{+0.89}_{-0.72}$	$0.40^{+0.38}_{-0.29}$

Table 6.3: Parameters and associated uncertainties for physical models of the dark matter mass structure, assuming a halo density profile as given in Equation 2, with (a) free ξ and r_s determined from the mass-concentration relation, and (b) free ξ and concentration c , and assuming (a) Chabrier and (b) Salpeter IMFs. A Chabrier IMF requires the presence of more dark matter centrally than the NFW+abundance matching prediction; conversely, a Salpeter IMF requires less dark matter than the NFW+abundance matching prediction.

population models from spectral absorption features would provide an independent probe of the IMF which would allow this degeneracy to be broken, and will be the topic of a future work. However, we also note that high-mass ETGs have been shown to require Salpeter-like IMFs in general (e.g. Auger et al., 2010b; van Dokkum & Conroy, 2010; Cappellari et al., 2012); though the EELs sources span a range in mass, we note that their masses are generally high, which here would imply the existence of sub-NFW haloes and the removal of dark matter from their central regions. We discuss this investigation further in Section 6.5.2.

6.5 Discussion

6.5.1 The fundamental plane

The rotation of the nominal fundamental plane from the virial prediction follows from the fact that the virial plane assumes that

$$(6.6) \quad \frac{M_{tot}}{L} = \text{constant}$$

across the ETG population, whereas, in reality,

$$(6.7) \quad \frac{M_{tot}}{L} = \frac{M_{tot}}{M_\star} \frac{M_\star}{L}$$

with both terms on the right-hand side of Equation 6.7 potentially dependent on other galaxy properties. In particular, the first term may vary across the plane if the dark or light mass structure is non-homologous, whilst the second term may vary due to non-homology in the stellar populations, such as the stellar age and IMF (though structural non-homology in the stellar mass may contribute to this term as well.) The EELs sources are massive,

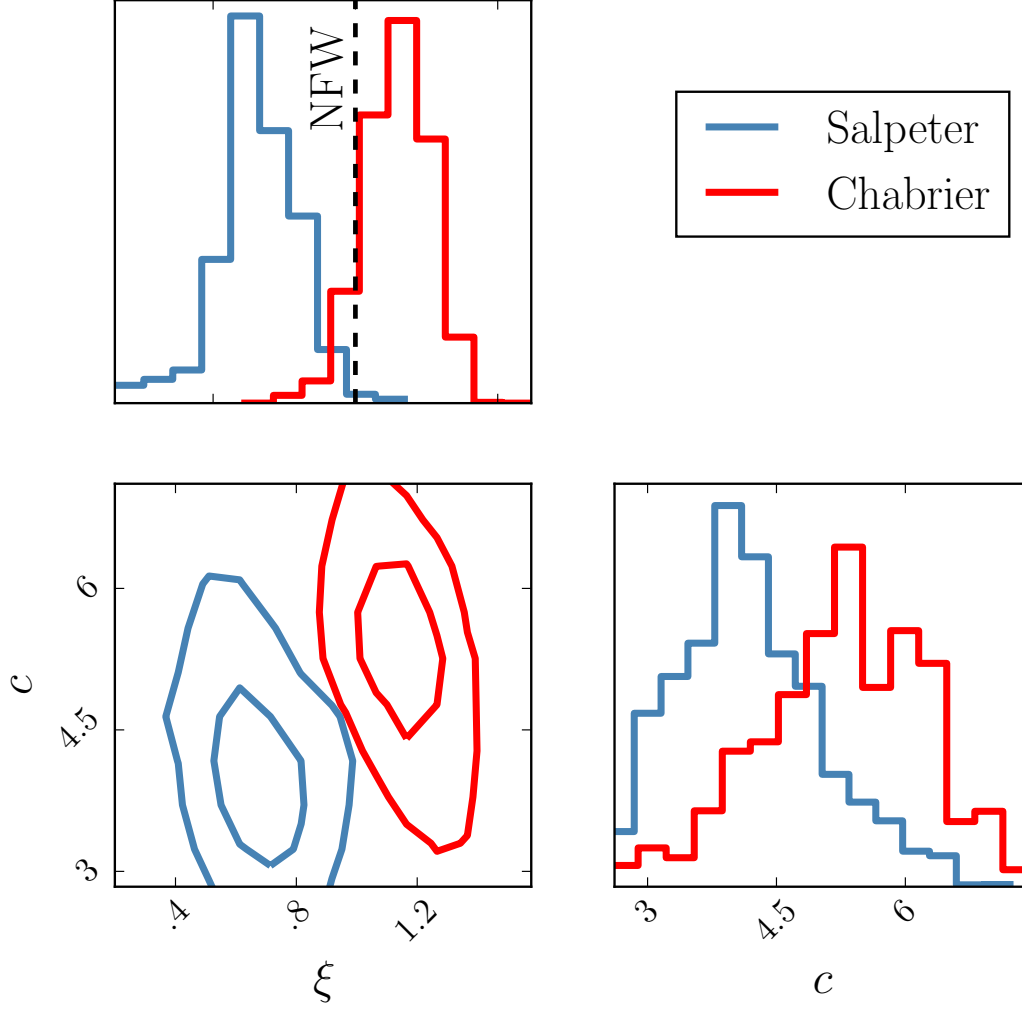


Figure 6.3: Inference on the structure of the dark halo for Chabrier (red) and Salpeter (blue) IMFs, compared to the expected inner slope from an NFW profile ($\xi = 1$; black dashed line). The measurement of the inner slope depends strongly on the form for the IMF that is assumed, with a bottom-heavy (Salpeter) IMF requiring less dark matter centrally than the NFW prediction and a bottom-light (Chabrier) IMF requiring more dark matter than the NFW prediction. An independent IMF probe is needed if we are to robustly distinguish between these scenarios; however, assuming that the EELs sources have Salpeter-like IMFs, consistent with previous IMF studies, implies the presence of sub-NFW haloes and points towards the action mergers and accretion in removing dark matter from the central regions.

compact ETGs, and their kinematics probe only their central regions, where the dark matter contribution to the mass is minimal. In this study, we have exploited this fact to attempt to disentangle these two contributions to the fundamental plane tilt, and thus to the behaviour of M_{tot}/L across the population.

In Section 6.3, we found the fundamental plane of the EELs to be oriented *in between* the nominal fundamental plane and the virial plane (and marginally consistent with both, i.e. at the 2σ level). Due to the low dark matter fractions in the centres of these systems (the kinematics imply masses within the effective radius on average 40% larger than the Chabrier stellar masses but 17% *smaller* than the Salpeter stellar masses, and previous studies indicate that Salpeter-like IMFs might be expected in these massive ETGs), this has implications for our understanding of ETGs as a population. Firstly, the rotation of the fundamental plane cannot be explained entirely by a mass-dependent dark matter fraction. If the rotation of the fundamental plane were solely due to mass-dependent dark matter fractions, then the low dark matter fractions of the EELs sources should cause their fundamental plane to be nearly aligned with the virial plane; that this is not the case implies that baryonic properties must also be mass-dependent. Conversely, if the fundamental plane tilt were solely due to mass-dependent baryonic properties, then the EELs fundamental plane should be aligned with the nominal fundamental plane. The fact that this is not the case either indicates, then, that *ETGs are non-homologous in both their dark and their luminous properties*. Previously, Bolton et al. (2008) showed that the fundamental plane tilt for the sample of ETGs spanning a similar mass range could be explained entirely by deviations from *structural* homology, as opposed to deviations from *dynamical* homology, though they were unable to disentangle the relative importance of the contributions from the dark and luminous mass components. Interpreting our result in the light of that study, we can update our conclusion above to say that ETGs are non-homologous in both their dark and their luminous *mass structures* and stellar populations.

We underline this result by further constructing the relation between mass, $\log M \propto 2\log\sigma + \log R_e$, and luminosity, $\log L \propto \log I_e + 2\log R_e$, in Figure 6.4. This projection of the fundamental plane is edge-on for both the virial plane and the nominal fundamental plane, such that $\log L \propto \log M$ corresponds to the virial plane (as follows from Equation 6.6) and $\log L \propto 0.75\log M$ is consistent with the local fundamental plane parameters (Jørgensen et al., 1995). We fit a linear relation between $\log M$ and $\log L$ following the methods of Section 6.3, and find a slope $\alpha_{LM} = 0.85 \pm 0.21$, which is also shown in Figure 6.4. As with the fundamental plane, this result places the $\log L - \log M$ relation in between the virial plane and nominal fundamental plane expectations, and seems to confirm the idea

that the low dark matter fractions of the EELs sources removes some, but not all, of the systematic trends in M_{tot}/L that occur across the ETG population. We note, however, that larger samples of lensed (or low-redshift) compact ETGs are needed in order to make more stronger claims in this regard.

We also note the difference in the fundamental plane of these compact ETGs from that of cluster ETGs at similar redshifts. In Chapter 4, we presented the fundamental plane of massive ETGs in the cluster MACSJ0717 at $z = 0.545$; this sample has the same mean velocity dispersion as the EELs sources ($\mu_\sigma \sim 200 \text{ kms}^{-1}$), but its fundamental plane is extremely consistent with that of local ETGs, with no evidence for a tilt (see Figure 6.2). The main implication of this result is that these cluster ETGs have already experienced the majority of their structural evolution by $z \sim 0.5$; contrary to the EELs sources, which have remained compact. In Chapter 5, we also showed that the EELs sources do not appear to occupy unusually dense environments. Taken together, this pair of results is interesting as it provides further evidence for the accelerated growth of galaxies in dense environments: if any massive compact galaxies existed in MACSJ0717, they would be some of the brightest ETGs in the cluster and so should have been included in that study. The fact that they are not suggests that all the initially-compact ETGs in that dense environment have already evolved into larger systems like those we see locally, whereas the EELs sources, in less dense environments, remain at an earlier stage of their evolution. Though a more quantitative statement would require a thorough characterisation of the selection functions of these two samples, this direct comparison strongly highlights the differences between these two populations.

Finally, we measure the fundamental plane to have a small scatter, $\log \sigma = -1.31 \pm 0.19$, comparable to that of the normal fundamental plane (Auger et al. 2010a find the SLACS lenses to have $\sigma = 0.049 \pm 0.009$). If scatter in the dark and the light structure contributed equally to the overall fundamental plane scatter, then we would expect to find a reduced intrinsic scatter in the fundamental plane of the EELs sources, where the low dark matter fractions should make the contribution to the scatter from the dark matter minimal. The fact that the intrinsic scatter is not decreased when variations in the dark matter are effectively removed in this way implies that this scatter is largely due to variations in the stellar structure and stellar populations, and that the dark matter structure in ETGs – at least in their central regions – is subject to less variation. Equally, though, our sample size is small, which makes statistical properties of the sample, such as the intrinsic scatter, subject to uncertainty.

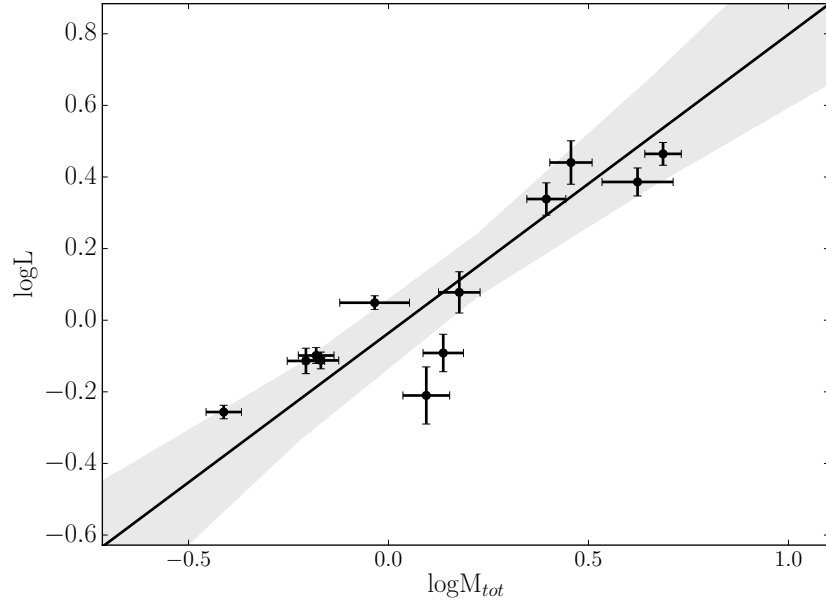


Figure 6.4: The mass-luminosity relation for the EELs sources, with $\log M_{tot} = 2\log\sigma + \log R_e - 5$ and $\log L = \log I_e + 2\log R_e - 10$ for σ in kms^{-1} , I_e in $\text{L}_\odot\text{kpc}^{-2}$ and R_e in kpc (note that, in each case, a constant is subtracted from the each axis to minimise covariance between the inferred intercept and slope). As with the fundamental plane, this relation for the EELs sources lies in between expectations based on the virial plane and the nominal fundamental plane, and suggests that the low dark matter fractions of the EELs sources removes some, but not all, of the systematic trends in M_{tot}/L that occur across the ETG population.

6.5.2 Physical models

In Section 6.4, we constructed toy models for the mass structure in the EELs sources in order to probe their dark matter distributions. Since we cannot currently constrain the IMFs in these systems, we constructed these models assuming a universal IMF which is either Chabrier and Salpeter, and found that the inference on the halo structure is (perhaps unsurprisingly) very sensitive to the choice of IMF. If a Chabrier-like IMF is assumed, we find the sources to require *more* dark matter in the centre than implied by the abundance matching+NFW prediction, whereas the use of a more bottom-heavy Salpeter IMF requires *less* dark matter centrally than the abundance matching+NFW prediction. These two results are very different and imply correspondingly different physics: whilst a sub-NFW dark halo would imply the action of either strong AGN feedback or merger processes in removing dark matter from the centre, a super-NFW halo would indicate an important role for initial adiabatic processes during galaxy assembly. It might be possible to distinguish

between these scenarios by using higher-signal-to-noise spectra to independently constrain the IMF in these systems via stellar population modelling, though the difficulty here would be in robustly measuring the continuum, given that the spectra contain the light from both the lens and the source galaxies. This is something we will explore in a future work.

For now, however, we combine a number of recent developments in our understanding of ETGs to sketch a feasible picture of the structure and evolution of the galaxies in our sample. At the outset, we emphasise our assumption that all the galaxies in our sample have the same IMF: in reality, their mass range may imply a variation in the IMF across the sample, which would complicate the following picture; nevertheless, given that the non-universality of the IMF remains not well understood and that we cannot constrain the IMF from our data, we adopt this as a reasonable first model.

First, existing evidence suggests that ETGs of comparable stellar mass to the EELs sources require bottom-heavy IMFs (e.g. Auger et al., 2010b; van Dokkum & Conroy, 2010; Cappellari et al., 2012), and further, that the IMF of ETGs may be most bottom-heavy in their central regions (Martín-Navarro et al., 2015a and Chapters 7 and 8). Seen in the context of the inside-out growth scenario, in which ETGs grow in two stages by the formation of a compact core at high redshifts, followed by the accretion of lower-mass systems at large radii to form an extended envelope, this suggests that the formation of the initial core may occur in fundamentally different star formation conditions from the lower-mass systems that it subsequently accretes. In this picture, the EELs sources, which seem to have experienced very little accretion so far, should have very bottom-heavy IMFs, comparable with the most central regions of local ETGs. As shown in Figure 6.3, this would then imply that their halo profiles are sub-NFW. What is interesting about this is that one of the most promising mechanisms for the removal of dark matter from the inner regions of ETG haloes is dynamical friction from infalling satellites during merger events, which has been shown in simulations to be effective at transferring energy from the infalling objects to the halo and therefore causing the latter to expand (Laporte et al., 2012).

In Chapter 5, we revealed evidence for ongoing merger activity in the *luminous* structure of these systems in the form of faint envelopes surrounding their compact cores; the possibility that they also have sub-NFW haloes may therefore present further evidence for their growth by mergers and accretion. This independent evidence, based on their *dark* structure, provides further insight into their evolution and presents new potential evidence for the importance of dry merging in the growth of ETGs.

6.6 Conclusions

We have presented and modelled spectra and the resulting stellar kinematics for 13 early-type/early-type lenses (EELs) to explore the nature of the fundamental plane and the dark and light mass structure of the source galaxies. Our main conclusions are as follows.

1. Relative to normal ETGs, the fundamental plane of the EELs sources is tilted towards, but still inconsistent with, the virial plane. Since the EELs sources are compact systems with small effective radii, their luminosity-weighted kinematics probe the very central regions where the dark matter fraction is low; this means their fundamental plane is sensitive predominantly to their *luminous* material, as opposed to their *dark matter*. The fact that their fundamental plane is rotated relative to both the fundamental plane of normal galaxies and the virial plane indicates that the properties of both the stellar populations and the dark matter are responsible for the well-known fundamental plane tilt, i.e. vary as a function of galaxy mass.
2. The intrinsic scatter of the fundamental plane of the EELs sources is small ($\log \sigma = -1.31 \pm 0.19$), but comparable to that of the fundamental plane of normal galaxies. This implies variations in the inner dark matter structure do not contribute significantly to the scatter, which must instead be driven by scatter in the stellar properties.
3. The halo structure can only be constrained, in the context of well-motivated models, when a universal IMF is assumed. However, on the basis of mounting evidence that massive, compact ETGs should have bottom-heavy IMFs, these galaxies appear to be best characterised by a dark matter halo with a sub-NFW inner density slope. This is consistent with a picture in which these galaxies are growing by successive mergers and accretion of smaller objects, and may therefore provide further evidence for their inside-out growth.

M87 HAS A RADIALY VARYING IMF: CONSTRAINTS FROM DYNAMICAL MASS MODELLING

Abstract

We present the first constraints on stellar mass-to-light ratio gradients in an ETG using multiple dynamical tracer populations to model the dark and luminous mass structure simultaneously. We combine the kinematics of the central starlight, two globular cluster populations and satellite galaxies in a Jeans analysis to obtain new constraints on M87's mass structure, employing a flexible mass model which allows for radial gradients in the stellar mass-to-light ratio, and find that we can rule out a constant stellar mass-to-light ratio with $> 99\%$ confidence, in the context of our models. Modelling the stellar mass-to-light ratio as following a power law with radius, we infer a power-law slope $\alpha = -0.54 \pm 0.01$ (statistical uncertainties), and use stellar population modelling of high-resolution 11-band HST photometry to show that such a steep gradient cannot be achieved by variations in only the metallicity, age, dust extinction and star formation history if the IMF remains fixed. On the other hand, combining our dynamical inference with the photometry, the stellar mass-to-light ratio gradient that we find is consistent with an IMF whose inner slope changes such that it is Salpeter-like in the central ~ 0.5 kpc and becomes Chabrier-like within the stellar effective radius. This adds to recent evidence that the non-universality of the IMF in ETGs may be confined to their core regions, and points towards a picture in which the stars in these central regions may have formed in fundamentally different physical conditions.

7.1 Introduction

The distribution of masses with which stars form is a fundamental property of a galaxy, and has an impact on virtually everything that we subsequently observe. However, the nature of the IMF in environments beyond our Milky Way remains uncertain. Whilst the IMF appears to be strikingly uniform across the diversity of environments within our own Galaxy (Bastian et al., 2010), and adequately described by a simple broken power law (Kroupa, 2001), recent years have brought to light an accumulation of evidence that the same may not be true extragalactically. Independent techniques based on strong gravitational lensing and stellar kinematics (e.g. Auger et al., 2010b; Cappellari et al., 2012) have indicated that more massive ETGs have more mass in stars than is predicted by a Milky-Way-like IMF, which the analysis of stellar-surface-gravity-sensitive spectral features has attributed to an excess of low-mass stars (van Dokkum & Conroy, 2010); this suggests a scenario in which the IMFs of more massive galaxies are more *bottom-heavy* than that of the Milky Way.

However, the astrophysical processes underlying these results remain extremely uncertain. The observed size evolution of ETGs (e.g. Daddi et al., 2005; van Dokkum et al., 2010) supports the idea that these systems grow significantly over time, primarily via minor mergers and accretion (Naab et al., 2009; Hopkins et al., 2009). If the IMF is non-universal, then the link between the formation conditions of the first stellar populations and the IMF at $z = 0$ is complicated by the fact that the IMFs of the stellar populations formed *in situ* and those that were accreted may differ. Moreover, the interpretation of the observed variations of the IMF as a function of galaxy velocity dispersion (Treu et al., 2010) is further complicated by (a) the degeneracy between dark and stellar mass, which so far has had to be broken by assuming a simple form for the halo (which was shown by Auger et al., 2010b to affect the strength of the correlation that is inferred), and (b) the fact that the calculation of global IMF mismatch parameters (see Equation 7.12) depends on luminosity-weighted properties integrated over some aperture (e.g. the Einstein radius of a lens, or a spectral aperture), introducing a non-uniformity between measurements and making it difficult to interpret trends across the galaxy population quantitatively.

Recently, a key step towards overcoming these limitations was provided by Martín-Navarro et al. (2015a), where gravity-sensitive spectral features were analysed as a function of radius for three nearby ETGs. The result showed that the two high-mass ETGs exhibit significant radial IMF gradients – with bottom-heavy IMFs in their central regions, which become Milky-Way-like at larger radii – whilst the IMF of the lower-mass system is consistent with being Milky-Way-like at all radii. In the context of the two-phase scenario

of ETG formation – in which a compact core forms at early times, followed by lower-density wings due to the accretion of lower-mass satellites (e.g. Naab et al., 2009) – this result points towards a picture in which the initial star formation processes in the progenitors of ETG cores are fundamentally different from those in lower-mass galaxies.

However, whilst stellar population studies such as this can suggest a radial dependence of the fraction of low-mass stars – and therefore the low-mass end of the IMF – they cannot formally provide any information about the high-mass end of the IMF. This must be investigated using probes such as dynamics and strong gravitational lensing, which are sensitive to the *total* stellar mass-to-light ratio Y_* , or the IMF *normalisation*, rather than the IMF *slope*. The first dynamical study of this kind, however – which used molecular gas kinematics to dynamically trace stellar mass-to-light ratio in the inner 1-2 kpc of seven massive ETGs – found a large scatter in both the overall IMF normalisation (ranging from sub-Chabrier to super-Salpeter) and the slope of the radial profile (including rising, falling and flat profiles), which furthermore did not seem to correlate with any global galaxy properties (Davis & McDermid, 2017).

One problem with focusing exclusively on *central* kinematics (i.e. gas or stars) is that the mass contribution from the dark matter halo cannot be well constrained, and must consequently be either subject to strong assumptions or ignored, which adds significant uncertainty to the resulting measurement of any stellar mass-to-light ratio gradient. To make progress, more extensive modelling must be carried out in order to infer the stellar mass and the dark halo structure simultaneously; however, these two mass components are extremely degenerate, and can only be robustly separated, in the context of a model, by combining multiple mass probes, each of which independently measures the gravitational potential. In Chapter 3, we combined multiple independent dynamical tracers of the potential in the giant elliptical M87, the BCG in Virgo, to infer the black hole mass, the structure of the dark matter halo and the stellar mass-to-light ratio – which was assumed to be uniform across the galaxy – and found that M87 could be best described by a centrally cored dark halo but that the inference on the IMF was dependent on our assumptions about the stellar orbital anisotropy. Here, we overcome this limitation using updated kinematic data and a more flexible model to extend the previous analysis to investigate the possibility of radial gradients in the stellar mass-to-light ratio for the first time.

The chapter is structured as follows. In Sections 7.2, 7.3 and 7.4 we introduce the data, our dynamical modelling and our statistical modelling, the main results of which we present in Section 7.5. Section 7.6 compares our dynamical constraints with expectations from stellar population modelling; we then discuss our findings in Section 7.7 and summarise in Section 7.8.

7.2 Data

To disentangle the contributions to M87's gravitational potential from the stellar mass and the dark matter, we require the kinematics of multiple independent tracer populations to simultaneously satisfy the Jeans equation for the same potential. Here, we use the kinematics of stars, globular clusters and satellite galaxies, which span a radius range from ~ 10 pc to 1 Mpc. In the following sections, we summarise the data used to characterise these different tracer populations.

7.2.1 Photometry

Use of the Jeans equation requires knowledge of the luminosity density $l(r)$ of each population from which kinematic tracers are drawn. This in turn depends on high-quality photometry. The datasets we use for this purpose are identical to those presented in Chapters 2 and 3, which should be referred to for further details. A summary is provided below and in Figure 7.1.

For the **stellar surface brightness**, we model the radial V-band profile of Kormendy et al. (2009) with a Nuker profile (exactly as in Chapter 3). Assuming spherical symmetry, we deproject this profile to give the 3D luminosity density shown (with uncertainties) in Figure 7.1.

For the **globular clusters**, we use the colour and radial profiles presented in Chapter 2, in which CFHT/MegaPrime imaging in the *ugriz* bands was used to compile a globular cluster candidate catalogue. We model the distributions of the two (red and blue) globular cluster populations, in addition to those of interloping objects. The globular clusters were treated probabilistically as being drawn from the red and blue populations, each of which is described by a Sérsic radial profile, a Gaussian luminosity function and radially-dependent Gaussian colour profiles. The 3D deprojected radial profiles for the red and blue populations are also shown in Figure 7.1.

Finally, for the **satellite galaxies**, we do not compile a photometric sample, as we would expect this to be incomplete and the selection function intractable; we therefore incorporate this population into our model using the robust mass estimator presented in Watkins et al. (2010), rather than a Jeans analysis. This is described in more detail in Chapter 3 and Section 7.3.

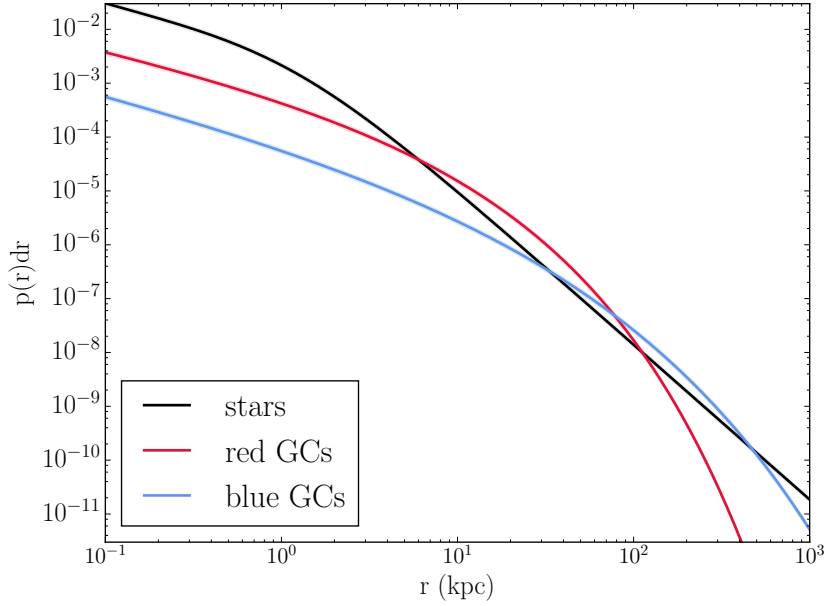


Figure 7.1: Normalised probability distribution functions for the deprojected 3D luminosity profiles of the dynamical tracer populations as a function of radius. Uncertainties are included, but are small. The stellar profile is modelled with a Nuker profile as in Equation 1; each globular cluster population follows a Sérsic profile in radius.

7.2.2 Kinematics

As our dynamical model is based on solutions to the spherical Jeans equation, we combine the photometric information of Section 7.2.1 with the kinematic data summarised below. The data used in this study are similar to those described in Chapter 3, but differ in two key respects. First, the globular cluster sample is almost doubled due to recent MMT/Hectospec observations (e.g. Ko et al., 2017) and samples M87’s surroundings more representatively and extends to larger radii. Second, the stellar kinematics from SAURON are superseded by new, higher-signal-to-noise kinematics from MUSE.

We combine **stellar kinematics** from two datasets which are complementary in spatial resolution and extent. In the central $2''$ (~ 170 pc), we use the velocity dispersions presented in Gebhardt et al. (2011), which were obtained using adaptive optics on Gemini/NIFS and binned radially in bins of $\log r = 0.16''$, with a spatial resolution of $\sim 0.1''$ ($= 80$ pc) and a signal-to-noise generally > 50 . At larger radii, we introduce a new measurement of the 2D velocity dispersion profile obtained using VLT/MUSE. The central arcminute of M87 was observed for one hour on the night of June 28 2014 during the MUSE science verification

phase, and the data are available in the ESO archive. We reduced the datacube using the standard ESOREX pipeline and modelled the resulting spectra, binned to 0.6 arcsecond pixels, in the rest wavelength range 5000-5775Å. We follow the methods of Chapter 4 by modelling the spectra as the linear combination of a set of stellar templates from the INDO-US library and an additive order-10 polynomial to account for the continuum. The velocity dispersion map that we obtain is consistent with that presented in Emsellem et al. (2014) and with the SAURON data at the 2σ level, though we note that our MUSE velocity dispersions rise more steeply in the central arcsecond, more closely following the NIFS kinematics in that region (this may in part be due to differences in the PSF), and is shown in Figure 7.3. For the dynamical modelling, we impose a minimum uncertainty of 5% to account for systematic uncertainties due to template choice.

For the **globular clusters**, we update the kinematic catalogue of Strader et al. (2011) to include the new globular cluster candidates that have been observed with MMT/Hectospec and are available on the CfA archive. The original Strader et al. (2011) catalogue combines measurements for 451 globular clusters – obtained using Keck/DEIMOS, Keck/LRIS and MMT/Hectospec – with literature data to provide line-of-sight velocity measurements for a total of 612 globular clusters within 240 kpc of M87. To supplement this, we cross-correlate the 2391 objects within ~ 1 degree of M87 with measured velocities from MMT/Hectospec with the photometric globular cluster catalogue of Chapter 2 and select as globular cluster candidates those objects which satisfy the colour cuts that were used in that photometric study. We expect some of these objects to be contaminant stars, but our probabilistic treatment of the globular cluster candidates accounts for this, as our photometric model assigns probabilities to each candidate of belonging to the red or blue globular cluster populations, a uniformly-distributed outlier population or a second outlier population following a model for the colour-magnitude distribution of the Sagittarius stream; during our dynamical inference, we stochastically sample from these probabilities at each step (see Section 7.4 for further details). We further impose a cut in line-of-sight velocity relative to M87 of 800 km s^{-1} and a cut in galactocentric radius of 0.9 degrees (to reduce contamination from intracluster globular clusters). This gives 439 additional globular cluster velocities, resulting in a sample of 1051 objects overall.

For the **satellite galaxies**, we use the same catalogue of 60 Virgo galaxies as in Chapter 3; the reader should refer to that chapter for further details.

7.3 Dynamical model

We construct models for M87's mass density and the anisotropy of each tracer population using the spherical Jeans equation.

7.3.1 Mass components

Our model for the total mass density of the galaxy $\rho(r)$ consists of a dark matter halo, a stellar mass component and a black hole:

$$(7.1) \quad \rho(r) = \rho_{DM}(r) + \rho_{\star}(r) + \rho_{BH}(r).$$

Guided by the results of Chapter 3, in which a number of different models for the dark halo were investigated, we use a generalised Navarro-Frenk-White (gNFW) profile to describe the dark matter halo

$$(7.2) \quad \rho_{DM}(r) = \frac{\rho_0}{4\pi} \left(\frac{r}{r_s} \right)^{-\gamma} \left(1 + \frac{r}{r_s} \right)^{\gamma-3}$$

where the scale radius r_s , inner slope γ and density scale ρ_0 are parameters to be inferred. In Chapter 3, we found that this profile was flexible enough to provide a good description of the data at relatively low computational cost.

The black hole is simply a point mass M_{BH} at the galaxy centre such that

$$(7.3) \quad \rho_{BH}(r) = \frac{M_{BH}}{4\pi r^2} \delta(r).$$

7.3.2 Stellar mass-to-light ratio

The key step forward in this study with respect to the models of Chapter 3 is that we now model the stellar mass density with a mass-to-light ratio Υ_{\star} that follows a radially varying profile. Thus, the projected surface mass density $\Sigma_{\star}(R)$ is related to the surface brightness distribution $I_{\star}(R)$ as

$$(7.4) \quad \Sigma_{\star}(R) = \Upsilon_{\star}(R) I_{\star}(R)$$

for projected radius R . It is important to point out at this stage that our separation of the dark and stellar mass components will be dependent on our choice of parameterisation of $\Upsilon_{\star}(R)$. To assess the sensitivity of our inference to the assumed form of $\Upsilon_{\star}(R)$, we consider two profiles. Firstly, we consider a **power-law (PL) profile** in which

$$(7.5) \quad \log \Upsilon_{\star}(R) = \log \Upsilon_{\star,1} + \mu \log R,$$

with power-law index μ , and $Y_{\star,1}$ representing the stellar mass-to-light ratio at a projected distance of 1kpc from the galaxy centre. In this model, however, the stellar mass becomes unphysically large in the very centre, which prevents us from making meaningful inference on the black hole mass; we therefore fix $M_{BH} = 6.6 \times 10^9 M_\odot$ to be consistent with the inferences of Gebhardt et al. (2011) and Chapter 3.

In reality, however, there exists some covariance between M_{BH} and $Y_\star(R)$ that is important to explore; we therefore consider a second **Salpeter-to-Chabrier (SC) model** in which Y_\star tends to a finite value centrally and becomes Chabrier-like at large radii:

$$(7.6) \quad Y_\star(R) = \alpha_s Y_{\star,s} + \frac{Y_{\star,ch} - \alpha_s Y_{\star,s}}{R^2 + R_M^2} R^2$$

for stellar mass-to-light ratios assuming Salpeter and Chabrier IMFs $Y_{\star,s}$, $Y_{\star,ch}$ inferred from photometry (see Section 7.6), mismatch parameter relative to a Salpeter IMF α_s and 2D scale radius R_M . Thus $Y_\star(R \rightarrow 0) = \alpha_s Y_{\star,s}$ and $Y_\star(R \rightarrow \infty) = Y_{\star,ch}$. As α_s can be either greater or less than unity, this model makes no assumptions about whether Y_\star rises or falls with radius, and the free scale radius R_M allows for the possibility that Y_\star is not Chabrier-like at any radii (i.e. R_M is allowed to become large). In this model, we allow the black hole mass to be a free parameter. We also explore generalisations of the SC model in which the index to which the projected radius is raised is a free parameter and the outer asymptotic stellar mass-to-light ratio is allowed to vary (i.e. $Y_{\star,ch}$ becomes $\alpha_{ch} Y_{\star,ch}$); however, we find that our constraints on these extra parameters are weak and that their inclusion does not significantly affect our results.

7.3.3 Anisotropy

To assess the robustness of our inference against different assumptions regarding the orbital anisotropy of the tracers, we also consider two different anisotropy models. Firstly, we consider an **isotropic** model, in which all tracers have zero anisotropy, $\beta = 0$, at all radii. This is the simplest assumption that can be made here, and may be physically reasonable if we consider all tracer populations to have been built up from multiple merging and accretion events. Secondly, we consider a more sophisticated **anisotropic** model in which the stars follow a scaled Osipkov-Meritt profile

$$(7.7) \quad \beta_\star = \beta_{\star,0} \frac{r^2}{r^2 + r_a^2}$$

where the scale radius r_a and asymptotic anisotropy $\beta_{\star,0}$ are parameters to be inferred. Since the globular cluster kinematics are sparser, we cannot constrain such a complex

model for their anisotropies, and instead treat them as having constant anisotropies $\beta_{r,b} = \text{constant}$, where the subscripts r, b refer to the red and blue populations respectively. Since these two populations are dynamically independent, their anisotropies are also treated as such.

7.3.4 Large-radius mass

As noted in Section 7.2.1, the underlying luminosity distribution of the satellite galaxies is not well understood; the catalogues from which our data are drawn were selected spectroscopically and are thus almost certainly subject to some unknown selection function. We therefore incorporate this population into our dynamical model using the mass estimator presented by Watkins et al. (2010), which is constructed to be robust against the use of simple approximations to a tracer population’s true density distribution, and which we calibrate using massive haloes from the MultiDark simulation (Prada et al., 2012). This procedure is presented in Chapter 3 and gives a constraint on the total mass of M87 within the projected radius of the outermost satellite galaxy:

$$(7.8) \quad \log\left(\frac{M(r < 985\text{kpc})}{M_\odot}\right) = 14.11 \pm 0.19$$

which allows us to normalise the mass profile appropriately at large radii.

7.4 Statistical model

We compare the stellar line-of-sight velocity dispersions calculated from the Jeans equation directly with the measured line-of-sight velocity dispersions, giving a contribution to the likelihood

$$(7.9) \quad \ln L_{\star,k} = -\frac{1}{2}\left(\frac{\sigma_k - \sigma_m}{\delta_k}\right)^2 - \frac{1}{2}\ln(2\pi\delta_{\sigma_k}^2)$$

for the k^{th} stellar velocity dispersion measurement. Here, σ_k and σ_m represent the observed and model velocity dispersions respectively, and the uncertainty δ_k is the quadratic sum of the measured uncertainty Δ_k and a regularisation term Δ_\star , which we infer to allow the datasets themselves to select their preferred relative weighting.

We compare the model globular cluster velocity dispersions with the observed globular cluster velocities by modelling the velocity distribution of each globular cluster population as a Gaussian with the dispersion calculated from the Jeans equation. We then assign probabilities of belonging to either the red or the blue population to each globular cluster

candidate, based on its velocity, colour, magnitude and spatial information, and stochastically sample these probabilities at each step in our MCMC exploration. At any step in this stochastic sampling, the likelihood contribution from the k^{th} globular cluster is

$$(7.10) \quad \ln L_{GC,k} = -\frac{1}{2} \frac{v_k^2}{\delta v_k^2 + \sigma_m^2} - \frac{1}{2} \ln(2\pi(\delta v_k^2 + \sigma_m^2))$$

for globular cluster velocity v_k , measurement uncertainty δv_k and model velocity dispersion σ_m , which is that of either the red or the blue population depending on the globular cluster's classification at that sampling step.

For the satellite galaxies, we compute the mass enclosed within the radius of the outermost object and compare it with the mass calculated from the mass estimator, giving a single contribution to the likelihood

$$(7.11) \quad \ln L_{sat} = -\frac{1}{2} \left(\frac{\log M_{sat} - \log M_m}{\delta \log M_{sat}} \right)^2 - \frac{1}{2} \ln(2\pi \delta \log M_{sat}^2)$$

for model mass $\log M_m$, and the mass and uncertainty calculated from the mass estimator, $\log M_{sat} = 14.11$, $\delta \log M_{sat} = 0.19$, measured within $R_{out} = 985$ kpc.

For each of the isotropic and anisotropic runs, we explore the parameter space using EMCEE.

7.5 Results

Our inferences on the structure of the dark halo and the stellar mass are shown in Figures E1 (isotropic, PL), E2 (anisotropic, PL), E3 (isotropic, SC) and E4 (anisotropic, SC) in Appendix E, and reported in Table 1. Our inference on the mass structure is similar in all four cases, and the halo structure we infer is also consistent with the result of Chapter 3, in which a constant Υ_\star was assumed (though the size of the central dark matter core is less certain). Since the resulting stellar mass and halo profiles are similar for all models, we select the anisotropic PL model as our default model due to its simpler and more intuitive form for Υ_\star ; the resulting mass profile is then shown in Figure 7.3. For the central halo structure, we find a weak inner slope $\gamma = 0.07_{-0.05}^{+0.12}$ and a scale radius $r_s = 40.70_{-6.26}^{+9.56}$, at large radii, the virial mass is $\log(M_{vir}/M_\odot) = 13.98_{-0.10}^{+0.12}$ and the virial radius $r_{vir} = 1092_{-79}^{+82}$ kpc. We note that this is consistent with our mass estimate from the satellite galaxy sample, though towards the lower end of that constraint; this may indicate shortcomings in our assumption that M87's halo resembles those in MultiDark. Nevertheless, given that we compress the satellite galaxy information down to a single data point, it also provides our weakest constraint, with our inference being mainly driven by the stellar and

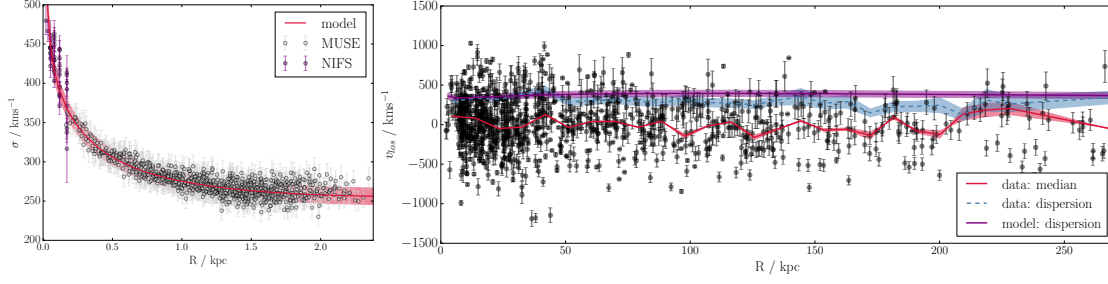


Figure 7.2: Comparison of model and data for M87’s kinematics, for both stars (left) and globular clusters (right). Note that our model does not explicitly fit the binned velocity dispersion of the globular cluster population.

globular cluster data. Indeed, we showed in Chapter 3 that removing this constraint from our likelihood calculations does not significantly change our inference. We are also able to well reproduce the kinematics of both the stars and the globular clusters, as shown in Figure 7.2.

The novel result from this modelling is that we can rule out a stellar mass-to-light ratio that is constant with radius with $>99\%$ confidence. We find $\Upsilon_{\star,1} = 4.45^{+0.07}_{-0.07}$, $\mu = -0.55^{+0.01}_{-0.01}$ (quantities defined in Equation 7.6), such that Υ_{\star} is a declining function of radius and becomes Chabrier-like by a radius of roughly 1 kpc. The stellar mass-to-light ratio profile that we infer is shown in Figure 7.4 (upper panel), with expectations from stellar population synthesis modelling overlaid (see Section 7.6). Our SC models agree with this result: for the anisotropic SC model, we find a high central stellar-mass-to-light ratio mismatch parameter $\alpha_s = 2.98 \pm 0.02$ and a small projected scale radius $R_M = 0.17 \pm 0.02 \text{ kpc}$, indicative of an IMF that is much heavier in the very central regions but consistent with being Chabrier-like across the rest of the galaxy. We confirm this last statement by running a model in which the outer asymptotic stellar mass-to-light ratio is also allowed to vary as $\Upsilon_{\star} = \alpha_{ch} \Upsilon_{\star,ch}$ and find that α_{ch} is consistent with unity.

Attributing this gradient wholly to a changing IMF indicates a relatively steep decline from a bottom-heavy Salpeter-like IMF in the central $\sim 0.5 \text{ kpc}$ to a Milky-Way-like Chabrier IMF at slightly larger radii (but still well within the stellar effective radius, and within the coverage of the MUSE kinematics). Figure 7.4 (lower panel) emphasises this decline by recasting the the stellar mass-to-light ratio in terms of the IMF mismatch parameter

$$(7.12) \quad \alpha_{chab} = \frac{\Upsilon_{\star,dyn}(R)}{\Upsilon_{\star,chab}(R)}$$

where $\Upsilon_{\star,dyn}(R)$ is the stellar mass-to-light ratio inferred from dynamics and $\Upsilon_{\star,chab}(R)$ is

that inferred from stellar population modelling assuming a Chabrier IMF. (Note that this shows the mismatch parameter α_{chab} at a *particular projected radius* R , as opposed to the mismatch that would be measured *within an aperture* of radius R .) Clearly, the mismatch parameter becomes large at the innermost radii; though $Y_{\star}(R \rightarrow 0)$ becomes unphysical in the PL model, the SC models also find $Y_{\star}(R \rightarrow 0) = \alpha_s Y_{\star,s}$ to be large (hitting the upper prior bound $\alpha_s = 3$; see Figures E3 and E4). Furthermore, all our models robustly predict a sharp decline of Y_{\star} within $R \sim 1\text{ kpc}$, well within the effective radius of the light. This is surprising as it implies that the IMF – and so, potentially, the star formation conditions – may differ significantly between the central regions – where the IMF is required to be heavier than the Salpeter parameterisation – and the rest of the galaxy. On the other hand, it forms a strikingly consistent picture with the recent work of Sarzi et al. (submitted), in which the same MUSE spectra were analysed using stellar population models to deduce a variation in the IMF slope from Salpeter-like to Chabrier-like within $0.4R_e$. We note that previous studies of M87's mass assuming a spatially uniform stellar-mass-to-light ratio have found a Salpeter-like IMF to be required (e.g. Murphy et al., 2011; Zhu et al., 2014; see also our isotropic model in Chapter 3); since these measurements are the luminosity-weighted integrals of $Y_{\star}(R)$ over apertures, they form a consistent picture with the results presented here, with the integral dominated by the high central Y_{\star} .

Of course, the stellar-mass-to-light ratio gradient that we infer may be due to gradients in the stellar age and metallicity in addition to those in the IMF; this is a possibility that we investigate in more detail in the following section.

7.6 Stellar population modelling

The key result of the dynamical modelling presented in this study is that the stellar mass-to-light ratio of M87 is a declining function of radius. The stellar-mass-to-light ratio as measured dynamically represents the summed contributions to both the mass and the light from across the stellar population(s) and so is sensitive to the age and metallicity of those populations, in addition to the integral over their mass function. From the dynamical inference alone, it is not possible to identify the driving factor behind the gradient that we infer.

To disentangle the contributions to the mass-to-light ratio due to these different stellar properties, we carry out stellar population modelling using the high-resolution eleven-band HST photometry of the central $17.5'' (= 1.4\text{ kpc})$ of M87 that was presented by Montes et al. (2014). These data span a wide range of filters from F336W (HST/WFPC2) to F850LP (HST/ACS), and are presented in that paper as surface brightness measurements within

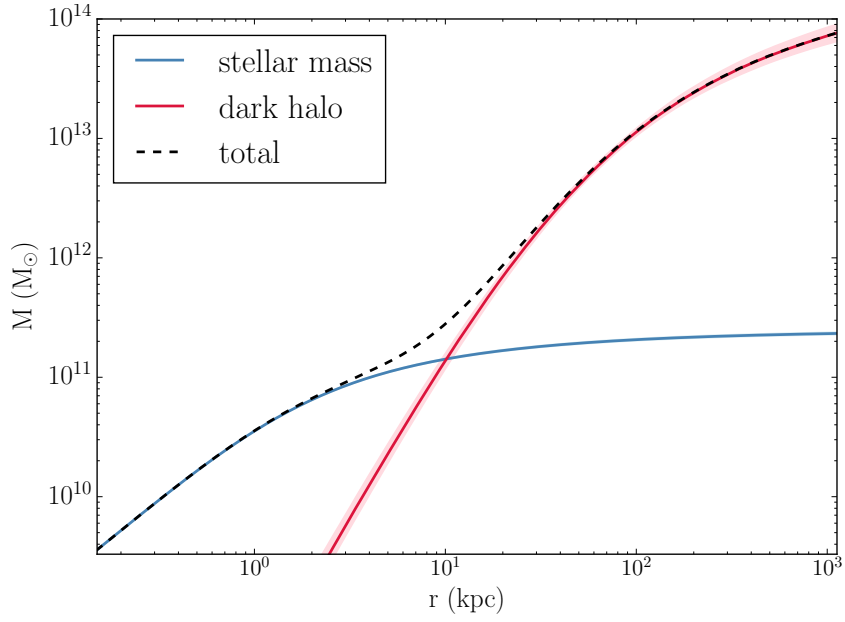


Figure 7.3: Inference on the dark, stellar and total mass profile in the anisotropic model. At radii ≤ 10 kpc $\sim R_e$, the stellar mass dominates, whereas beyond this, the dark halo becomes the main contributor to the potential. Our kinematic data extend from ~ 10 pc to 1 Mpc, which is the radius range spanned in this Figure.

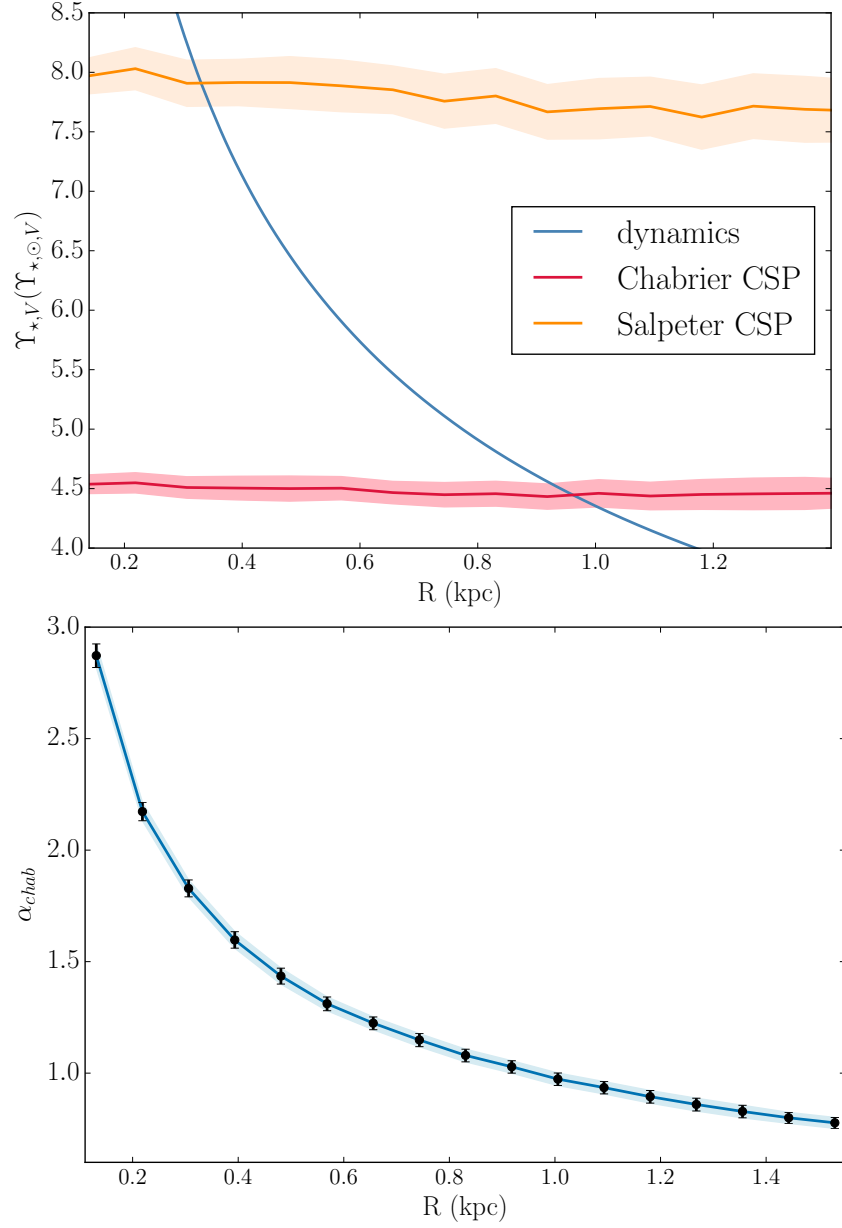


Figure 7.4: Top: The stellar mass-to-light ratio inferred dynamically, shown in blue, declines much more rapidly than can be achieved by gradients in the age, metallicity, star formation history and dust extinction of the stellar populations under a fixed IMF, suggesting that an IMF gradient may be the driving factor. Indeed, the stellar mass-to-light ratio is consistent with stellar population models that assume a Salpeter IMF at small radii, but consistent with stellar population models assuming a Chabrier-like IMF at larger radii. Bottom: the IMF mismatch-parameter α_{chab} as defined in Equation 15, again showing that the mismatch between the dynamically-inferred stellar mass-to-light ratio and the stellar population-modelling-inferred stellar mass-to-light ratio under the assumption of a Chabrier IMF increases towards the centre of the galaxy.

MODEL 1: $\Upsilon_{\star}(R) = \alpha \Upsilon_{\star,s} + \frac{\Upsilon_{\star,ch} - \alpha \Upsilon_{\star,s}}{R^2 + R_M^2} R^2$										
model	α	R_M	$\log \rho_0$	r_s	γ	β_{\star}	r_a	β_r	β_b	$\log M_{BH}$
isotropic	$2.84^{+0.05}_{-0.10}$	$0.16^{+0.02}_{-0.02}$	$8.96^{+0.07}_{-0.13}$	$29.78^{+4.50}_{-3.05}$	< 0.27	—	—	—	—	$7.5^{+0.20}_{-0.19}$
anisotropic	$2.98^{+0.02}_{-0.02}$	$0.17^{+0.02}_{-0.02}$	$8.54^{+0.09}_{-0.11}$	$55.04^{+10.10}_{-7.91}$	< 0.14	$0.92^{+0.04}_{-0.04}$	$25.16^{+3.54}_{-3.20}$	$0.31^{+0.20}_{-0.23}$	$0.27^{+0.16}_{0.17}$	$5.14^{+0.12}_{-0.11}$
MODEL 2: $\Upsilon_{\star}(R) = \Upsilon_{\star,1} R^{-\mu}$										
model	$\Upsilon_{\star,1}$	μ	$\log \rho_0$	r_s	γ	β_{\star}	r_a	β_r	β_b	
isotropic	$4.45^{+0.07}_{-0.07}$	$-0.55^{+0.01}_{-0.01}$	$9.47^{+0.03}_{-0.05}$	$16.87^{+0.94}_{-0.77}$	< 0.14	—	—	—	—	
anisotropic	$4.34^{+0.06}_{-0.06}$	$-0.54^{+0.01}_{-0.01}$	$8.76^{+0.12}_{-0.17}$	$40.70^{+9.56}_{-6.27}$	$0.07^{+0.12}_{-0.05}$	$0.93^{+0.04}_{-0.04}$	$15.47^{+2.48}_{2.06}$	$0.30^{+0.20}_{-0.24}$	$0.22^{0.16}_{-0.19}$	

Table 7.1: Final inference on M87’s mass profile for both the isotropic and anisotropic models, with a default M/L profile. We report the maximum-posterior values of our samples, along with the 16th and 84th percentiles as a measure of our uncertainty. Where γ hits the lower bound of the prior, we give the 95th confidence level. All quantities are measured in units of solar mass, solar luminosity, kilometres per second and kiloparsecs.

circular annuli of width $1''$. We can therefore use stellar population models to infer the age, metallicity, star formation history, dust extinction and stellar mass-to-light ratio as a function of projected radius, making this the ideal dataset to compare with our dynamical inference.

We model the photometry following the methods of Chapter 5. We use the composite BC03 stellar population models to compute apparent magnitudes in the 11 filters on a grid of stellar age T , metallicity Z , dust extinction τ_V and time constant τ of an exponentially decaying star formation history, and construct a spline interpolation model which allows magnitudes to be evaluated at any point within the grid; these magnitudes can then be scaled by the stellar mass. We then explore the posterior probability distribution of these parameters, along with the stellar mass, using EMCEE. We treat each radius ‘bin’ as completely independent, such that Z , T , τ , τ_V and M_\star can vary freely as a function of radius, and impose a ‘global’ IMF that is either Chabrier or Salpeter.

The resulting stellar-mass-to-light ratio profiles that we infer therefore tell us not only the difference in magnitude of the stellar mass-to-light ratio under different IMF assumptions, but also the stellar mass-to-light ratio slope that can be achieved by allowing gradients in all parameters *except* for the IMF. Figure E5 shows an example of our inference on the various stellar population properties for a Chabrier IMF: the gradients in metallicity and age are small, resulting in a significantly shallower stellar mass-to-light ratio gradient than that inferred dynamically (Figure 7.4); the result for a Salpeter IMF is qualitatively the same. It appears, then, that age, metallicity and star formation history variations cannot be driving the Y_\star slope, and radial gradients in some other property must be responsible for this. A radially varying IMF, falling from a Salpeter-like function to a more Chabrier-like one with increasing radius, may instead be the driving factor.

Based on these findings, we attempt to *infer* the IMF slope as a function of radius by constructing stellar population models which are required to fit the photometry and our inference on the projected mass profile simultaneously. We use the Flexible Stellar Population Synthesis (FSPS) models of Conroy et al. (2009); Conroy & Gunn (2010), which allow significantly more freedom in the form of the IMF than the BC03 machinery. We compute magnitudes on a grid of metallicities, ages and low-mass IMF slopes Γ , where for the IMF we assume a double power-law with the form

$$(7.13) \quad \frac{d \log N}{d \log m} \propto m^{-\xi}, \quad \begin{cases} \xi = \Gamma & m < 1M_\odot \\ \xi = 2.3 & m > 1M_\odot \end{cases}$$

such that the IMF follows the canonical (e.g. Chabrier, Kroupa, Salpeter) form at high masses, but is flexible at the low-mass end to allow the data to choose between bottom-heaviness ($\Gamma > 2.3$) and bottom-lightness ($\Gamma < 2.3$). Note that $\Gamma = 2.3$ corresponds to a Salpeter IMF and $\Gamma = 1.3$ corresponds to an IMF which is Milky-Way-like. Guided by the BC03 analysis, in which τ_V and τ change negligibly with radius (and τ is short), we consider SSPs with a fixed dust extinction parameterised as in Charlot & Fall (2000, similarly to

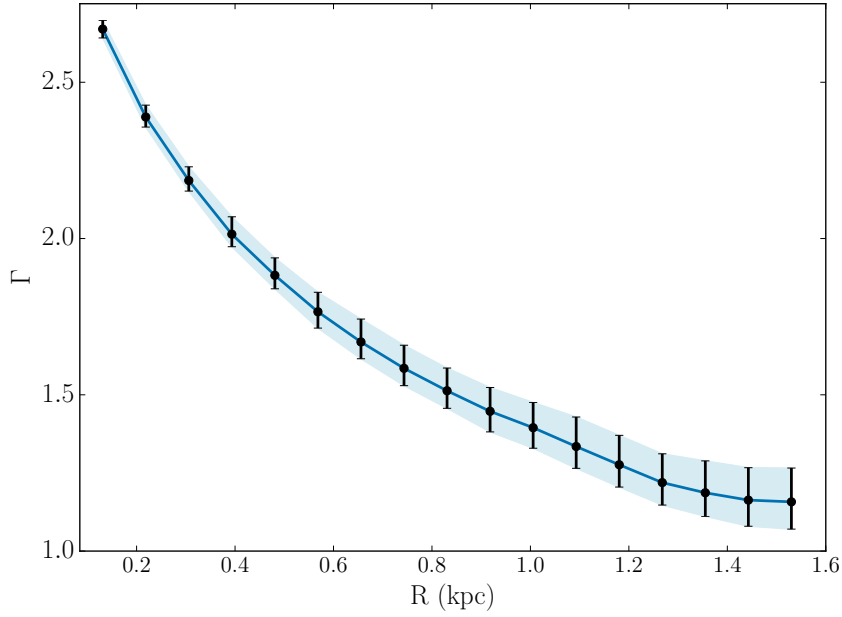


Figure 7.5: Inference on the slope of a broken-power-law IMF with slope $\xi = 2.3$ for $M > 1M_{\odot}$ and $\xi = \Gamma$ for $M < 1M_{\odot}$. We use FSPS to calculate magnitudes on a grid of ages, metallicities and Γ , and obtain the posteriors on these quantities based on 11-band HST photometry and our dynamical inference on the projected stellar mass as a function of radius. Our model clearly requires an IMF that becomes increasingly bottom-heavy towards the centre. Nevertheless, it is possible that alternative models which allow greater flexibility in other stellar population properties may also be able to reproduce the photometry and mass inference simultaneously.

BC03), and allow Z , T and Γ to vary freely with radius. We then require our model to reproduce the photometry under the condition that the mass profile follows that which we have inferred dynamically. Figure 7.5 shows our inference on Γ as a function of radius: we find that the IMF slope implied by the photometry and dynamics is super-Salpeter in the innermost radial bins and becomes approximately Milky-Way-like by the outermost bin. We find universally old stellar ages $\log T \sim 10.07$ Gyr and consistently supersolar metallicities; both consistent with our BC03 analysis. Of course, this model represents one possible explanation for M87’s stellar mass-to-light ratio gradient, but we cannot rule out the possibility that more flexible stellar population modelling, allowing freedom in a greater number of parameters, could also explain our result.

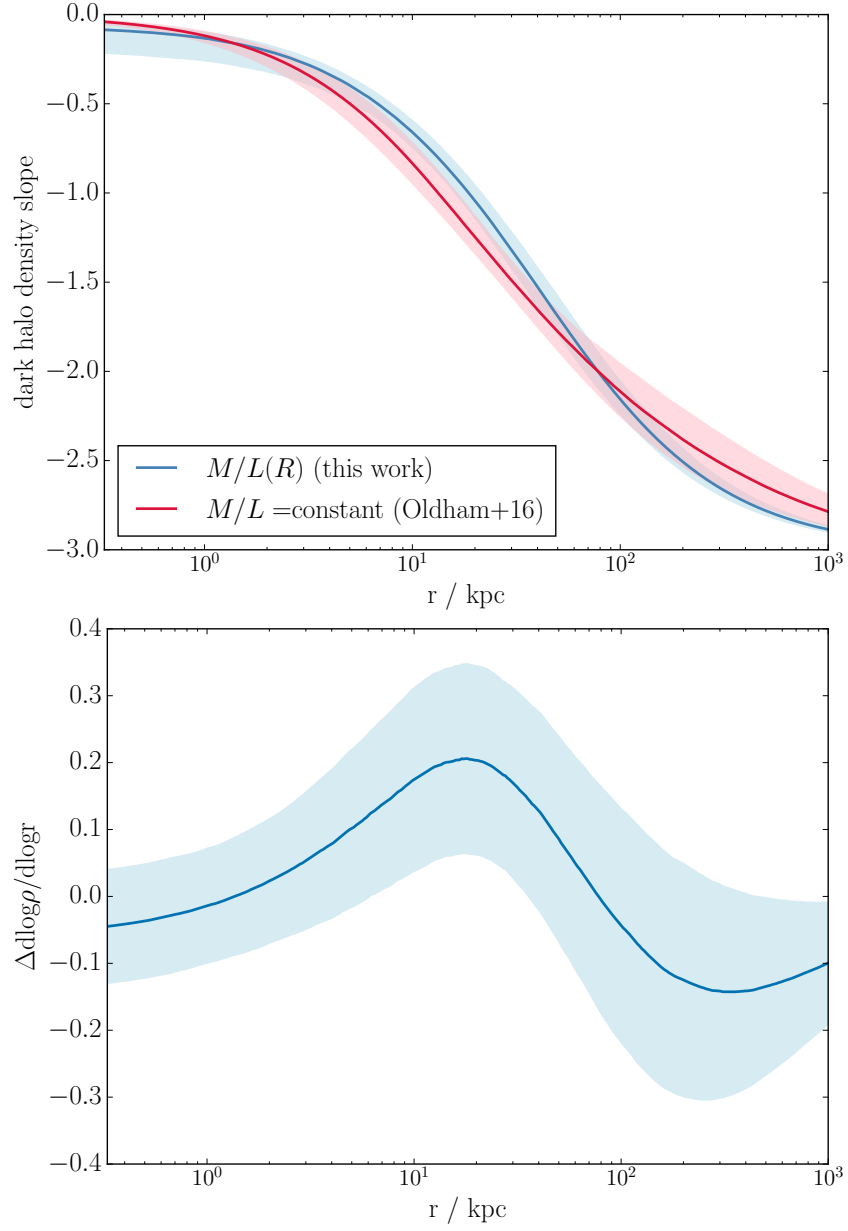


Figure 7.6: Top: Inference on the dark halo density slope as a function of 3D radius in this work (blue) and the work of Chapter 3 (red), in which a constant stellar mass-to-light ratio was assumed. The more flexible model used in this work leads to a slightly larger uncertainty on the halo structure at the smallest radii, and the slope is also slightly softer in this case. Bottom: Difference in slope between this work and Chapter 3 (i.e. blue curve minus red curve from the upper panel) as a function of radius: the differences vary distinctly with radius, but are nevertheless small, indicating that our inference on the halo structure does not depend crucially on the complexity of the stellar mass model.

7.7 Discussion

By dynamically modelling the stars, globular clusters and satellite galaxies in M87's gravitational potential, we have disentangled the contributions from the dark and stellar mass to show that (a) the dark halo is centrally cored, contrary to the predictions of dark-matter-only simulations, and (b) the stellar IMF is a declining function of radius. In the following sections, we discuss these two results in more detail.

7.7.1 Inference on the halo structure is robust

This study serves as an extension of the models of Chapter 3, which set the first constraints on M87's dark matter core. There, we investigated a number of different parameterisations for both the halo and the anisotropy profiles, and showed that the inference on the halo structure was robust against reasonable (and computationally feasible) changes in the model. However, we did not investigate the effect of our assumption of a constant stellar mass-to-light ratio, and this remained a dominant source of systematic uncertainty in our result. In this work, we have removed this source of uncertainty by showing that allowing stellar mass-to-light ratio gradients, whilst affecting the stellar mass profile and the finer details of the halo profile, does not remove the need for a central dark matter core. In Figure 7.6, we show a comparison between the inferred dark halo density slope from Chapter 3 and the present work. The core we find here is milder, and less well constrained in the centre, due to the increased flexibility of the model in that region. As shown in Figure 7.6 (bottom), we predict a steeper slope $\Delta\gamma \sim -0.1$ at the inner- and outer-most radii, and a shallower slope $\Delta\gamma \leq 0.2$ at intermediate radii: of course, some of this is a consequence of the analytic profiles that we impose for both the dark matter and the stellar mass, which limit the freedom that the slope is allowed at different radii. Nevertheless, this shows that the qualitative nature of the halo structure is robust against variations in the stellar mass-to-light ratio model; M87 does indeed appear to have less dark matter centrally than the NFW prediction, indicating the action of baryonic physics in the form of AGN feedback or dynamical friction during satellite infall (e.g. Laporte et al., 2012; Martizzi et al., 2012). The physical implications of this are discussed more extensively in Chapter 3.

7.7.2 Radial gradients in the IMF?

This study represents the first use of dynamical tracers to disentangle the mass contributions from a dark matter halo and a stellar component with a mass-to-light ratio gradient.

We implement two independent models for the stellar mass-to-light ratio, both of which robustly prefer a stellar mass-to-light ratio which declines strongly within the central ~ 1 kpc. Stellar population modelling of high-resolution multiband photometry suggests that this gradient is much stronger than can be achieved by gradients in either metallicity, age, extinction or star formation history, and therefore that a varying IMF – transitioning from a heavy Salpeter-like IMF in the very centre to a lighter Milky-Way-like IMF within a stellar effective radius – may be responsible. In Martín-Navarro et al. (2015a), stellar population modelling was used to measure radial gradients in the IMF slope in two high-mass ETGs – treating the IMF as a single power law which tapers off to a constant value below $M < 0.6M_{\odot}$ – and found that the slope at the effective radius was around half its value in the centre, implying a significant excess of low-mass stars in the central regions. Since then, van Dokkum et al. (2016) have found evidence for qualitatively similar trends in six other massive ETGs, which appear to be generally well-described as having centrally bottom-heavy IMFs which become Milky-Way-like beyond the central $\sim 0.4R_e$. Whilst the dynamical modelling presented in this study is sensitive to a different aspect of the IMF – that is, the integral over the IMF, rather than the fraction of low-mass stars – the radially-declining stellar mass-to-light ratio that we find can be naturally explained in terms of a decreasing fraction of low-mass stars. These independent and consistent results from dynamics and stellar population modelling therefore make a compelling case for the existence of a fundamentally different, bottom-heavy IMF in the innermost regions of massive ETGs.

This dependence of the form of the IMF on galaxy radius has important implications for our understanding of ETG assembly and evolution. First, that all but the very innermost regions have IMFs which are consistent with being Milky-Way-like suggests that the star formation conditions in these regions do not differ significantly from those found in lower-mass galaxies. In light of the recent accumulating evidence for the build-up of massive ETGs being dominated by repeated minor mergers and accretion (e.g. van Dokkum et al., 2010), this may be the result of a large amount of the larger-radius stellar material being originally formed in lower-mass Milky-Way-like systems which were subsequently accreted. On the other hand, the bottom-heavy IMFs in the cores of these ETGs imply fundamentally different star formation conditions in these regions; as these galaxies most likely form in the centres of the most massive dark haloes with the deepest potential wells, these different conditions may arise from the especially dense environments in which initial star formation must take place. In this picture – in which the compact central core forms at early times with a bottom-heavy IMF, followed by the ongoing accretion of lower-mass satellites with less bottom-heavy IMFs – the gradient that we infer here arises

naturally, and may also be expected to be particularly pronounced in M87, which sits at the high-mass end of the galaxy mass function, in the centre of a cluster, and surrounded by a vast population of globular clusters. An important future step in understanding the origin of these IMF gradients will be to examine their strength across the ETG mass function.

7.7.3 The importance of accounting for IMF gradients

M87 is a massive, nearby BCG for which extensive data are available, such that it is possible for us to construct models which constrain the gradient in its stellar mass-to-light ratio. However, for the majority of ETGs for which stellar mass-to-light ratios are measured, limitations in the data may make obtaining such constraints difficult. It is therefore important to consider the systematic uncertainties that are introduced by modelling a varying stellar mass-to-light ratio as constant. Given the previous work on M87 presented in Chapter 3, we are now in a position to do this.

First, we note from Section 7.7.1 that variations in the stellar mass model do not significantly affect our inference on the halo structure. This is encouraging as it further justifies previous attempts to disentangle the dark and luminous mass in ETGs in order to infer their halo structure (e.g. Newman et al., 2015). Nevertheless, M87 has a very low central dark matter fraction due to the cored nature of its halo, and it is possible that this may minimise the degeneracy between the dark and luminous mass. ETGs in more isolated environments than M87 may have significantly cuspiest haloes (Sonnenfeld et al., 2012, and Chapter 8); it is possible that the existence of stellar mass-to-light ratio gradients in these systems may have a greater impact on the inference on the halo structure. This is an issue that will need to be addressed in the future; the recently-discovered low-redshift lenses (Smith et al., 2015), which reside in a range of environments and for which both lensing and extensive dynamical information is accessible, may represent suitable opportunities for this.

In terms of the stellar mass, however, we note that our difficulty in breaking the mass-anisotropy degeneracy in the stellar mass in our constant- Y_* models – with the stellar mass changing significantly depending on the complexity of our anisotropy model – may be a result of the inadequacy of the constant- Y_* model. Indeed, when we implement this more flexible radially-varying Y_* model, our inference on the mass structure of both the dark and stellar mass agree regardless of our assumptions about the anisotropy. The implication is that care must be taken in interpreting the stellar mass that is inferred when a constant stellar mass-to-light ratio is assumed, and attention paid to which particular aspects of the data are driving the inference. Furthermore, we note that the isotropic model of Chapter 3

required a significantly larger stellar mass-to-light ratio. The role played by this source of systematic uncertainty in the correlations of the IMF mismatch parameter with other galaxy properties (such as the stellar velocity dispersion; see e.g. Auger et al., 2010b) also needs to be better understood. Finally, we emphasise that the existence of IMF gradients complicates the comparison of measurements in different galaxies based on data extracted over different physical apertures, and that care must be taken to ensure that meaningful comparisons are made.

7.8 Summary and conclusions

We have combined multiple kinematic tracers of the mass in M87 to disentangle the dark and luminous mass components and the stellar anisotropy, and have inferred the presence of stellar mass-to-light ratio gradients in this massive galaxy for the first time. Our main conclusions are summarised below.

1. The stellar mass-to-light ratio $Y_{\star,v}$ in M87 is a declining function of radius. Parameterising $Y_{\star,v}$ in the V-band as a power law, we find a slope $\alpha = -0.54 \pm 0.01$ and $Y_{\star,v} = 4.34 \pm 0.06$ at 1kpc from the centre.
2. Multi-band, high-resolution photometry indicates that such a strong stellar mass-to-light ratio gradient cannot be achieved by varying only the metallicity, age, dust extinction and star formation history of the stellar population if the IMF remains fixed. On the other hand, the stellar mass-to-light ratio gradient that we infer is consistent with M87 having an IMF which is Salpeter-like in the central ~ 0.5 kpc and becomes Chabrier-like at ~ 1 kpc.
3. Inference on the halo structure does not change significantly depending on whether or not these stellar mass-to-light ratio gradients are allowed for in the model, indicating that the separation of the dark and luminous mass is robust. However, moving forward, it will be important to account for the presence of stellar mass-to-light ratio gradients when interpreting stellar mass measurements and their implications for the non-universality of the IMF in ETGs.

DARK MATTER CONTRACTION AND VARYING STELLAR MASS-TO-LIGHT RATIOS IN 12 STRONG LENSES

Abstract

We present models of the dark and luminous mass structure of 12 strong lensing ETGs. We combine pixel-based modelling of multiband HST/ACS imaging with Jeans modelling of kinematics obtained from Keck/ESI spectra to disentangle the dark and luminous contributions to the mass. Assuming a gNFW profile for the halo and a spatially constant stellar-mass-to-light ratio Υ_\star for the baryonic mass, we infer distributions for Υ_\star consistent with IMFs that are heavier than the Milky Way's, but halo inner density slopes which are extremely cusped. Relaxing the assumption of a spatially constant Υ_\star such that it follows a power law with radius, we find that the majority of our sample has declining stellar-mass-to-light ratios (with a global power-law index $\alpha_{ML} = -0.22 \pm 0.08$) – consistent with IMFs which become less bottom-heavy with increasing radius – and halo inner slopes which are cuspier than the NFW prediction (with global inner slope $\gamma(R < R_e) = 1.66 \pm 0.11$) and consistent with theoretical expectations from adiabatic contraction models. This seems to confirm earlier evidence that the deviation of the IMF in massive ETGs may be confined to their central regions and that the relative importance of different baryonic processes depends on halo environment. We suggest a coherent explanation of these results in the context of the inside-out galaxy growth.

8.1 Introduction

Dark-matter-only simulations of a Λ CDM Universe predict that the dark matter haloes of all galaxies should look nearly self-similar, regardless of their mass scale, with a density profile which declines as r^{-3} at large radii and has a central r^{-1} cusp (Navarro et al., 1996, 2010). However, the dominance of baryonic matter in the centres of massive ETGs may significantly modify the dark matter structure from this simple expectation, such that the halo may become either centrally contracted due to the initial infall of gas (Blumenthal et al., 1986; Gnedin et al., 2004) or expanded due to dynamical heating during satellite infall (El-Zant et al., 2001; Nipoti et al., 2004; Laporte et al., 2012) or AGN-driven outflows (Governato et al., 2012; Martizzi et al., 2012). The halo structure is therefore key to determining the relative importance of different physical processes in ETG evolution.

However, since most gravitational probes are only sensitive to the *total* galaxy mass, determining the halo structure relies on robustly decomposing the mass into its dark and luminous contributions. This problem is complicated by the fact that the form of the IMF in ETGs is not well understood. Whilst the diversity of environments within the Milky Way are consistent with having one same ‘universal’ IMF (Bastian et al., 2010), independent evidence from lensing, dynamics and stellar population modelling suggests that the IMF of massive ETGs may be bottom-heavy, i.e. may have an excess of low-mass stars relative to the Milky Way (Auger et al., 2010b; van Dokkum & Conroy, 2010; Cappellari et al., 2012). Moreover, it is not yet clear whether these trends are driven by galaxy mass or metallicity (Treu et al., 2010; Cappellari et al., 2013b; Martín-Navarro et al., 2015b), and recent evidence from stellar population modelling of optical spectra has suggested that the IMF may also vary radially *within individual galaxies*, (Martín-Navarro et al., 2015a; van Dokkum et al., 2016) – with the deviations from a Milky-Way-like IMF confined to the central regions – though constraints on the IMF using molecular gas kinematics and stellar population modelling at near-infrared wavelengths have not confirmed this (Zieleniewski et al., 2017; Vaughan et al., 2016; Alton et al., 2017; Davis & McDermid, 2017).

Knowledge of the IMF in massive ETGs is essential for probing halo structure; however, it is also critical for understanding ETG assembly and evolution, since the IMF is a key diagnostic of the physical conditions in which a galaxy first formed stars. Furthermore, the measurement of stellar population properties in external galaxies depends on knowledge of the IMF, such that the assumption of an incorrect IMF may lead to systematic bias in properties such as metallicity and age. Methods to simultaneously determine the dark and luminous mass structure of ETGs are therefore extremely important.

Recently, progress has been made by combining mass probes across different spatial

scales such that, in the context of a well-motivated model, this dark/light degeneracy can be broken. Due to the need for multiple mass tracers, these studies have focused on massive systems in dense environments, where populations of satellite galaxies, globular clusters and planetary nebulae can be used to dynamically trace the mass out to the virial radius and so complement stellar kinematic and strong lensing constraints in the central regions. In BCGs, these studies have found haloes that are expanded relative to the NFW prediction (Newman et al., 2013, and Chapters 3 and 7). On the other hand, the single study so far of group-scale ETGs has found their haloes to be mildly contracted (and consistent with the NFW model within 2σ ; Newman et al., 2015). On the scale of isolated field ellipticals, breaking the dark/light degeneracy is more difficult due to the absence of large-radius tracers and the smaller physical size of the Einstein radius, but one study of a rare double source plane lens has found evidence for strong contraction on these scales (Sonnenfeld et al., 2012; see also Grillo, 2012, which obtained similar constraints by combining aperture mass measurements for an ensemble of lenses assuming a fixed IMF, and Sonnenfeld et al., 2015, which made no IMF assumptions and thus much weaker inference on the halo slope). These differences in the halo structure as a function of environment may represent a real trend in the relative importance of baryonic processes; however, the small number of such studies so far – especially of isolated systems – makes it impossible to draw meaningful conclusions. Improved techniques for extracting information from isolated strong lenses are therefore needed.

This study is a first attempt to address this need. So far, studies of the halo structure of isolated strong lenses have condensed the lensing information down to a single measure of the Einstein radius or a set of conjugate points whose positions must be focused in the source plane (Grillo, 2012; Sonnenfeld et al., 2012, 2015); however, the extended arcs in galaxy-galaxy lenses contain much more information than this, such that a full reconstruction of the lensed images allows a much more precise inference on the halo. In this study, we construct consistent dark and light mass models of 12 isolated strong lenses at the pixel level in order to explore their dark and luminous mass structure in detail.

The chapter is organised as follows. In Section 8.2, we introduce the data, the mass models and our lensing and dynamical modelling methods. Section 8.3 presents our main results, which we discuss in Section 8.4. Finally, we summarise and conclude in Section 8.5.

8.2 Data and modelling

We construct simultaneous strong lensing and dynamical models for 12 of the EELs presented in Chapter 5 (note that we exclude J1619, whose lensed features we were unable

to fully reproduce using a parametric source model). To this end, we combine imaging and kinematic data as detailed below.

8.2.1 Data

The EELs were observed using HST/ACS (GO: 13661, PI: Auger) as presented in Chapter 5. Each EEL was observed at two dither positions for ~ 500 seconds per dither in the F555W/F606W (henceforth *V*-band) and F814W (henceforth *I*-band) filters – with the former dependent on the lens redshift, and chosen to straddle the Balmer break – and reduced using ASTRODRIZZLE with a pixel scale of $0.05 \text{ arcsec pix}^{-1}$.

The EELs were also observed with ESI/Keck (Sheinis et al., 2002), and the spectra are presented in Chapter 6. Each EEL was observed for ~ 1 hour using a 0.75 arcsec slit; the spectra were extracted using a custom-made PYTHON code and the first and second velocity moments inferred by modelling the lens and source components simultaneously using stellar templates from the INDO-US library, as described in Chapter 6. For this study, we extract kinematics for the lens galaxies over rectangular apertures extending 1 arcsec either side of the lens such that they probe the mass beyond the Einstein radius (which is $\sim 0.5 \text{ arcsec}$ for a typical EEL; see Table 5.2 in Chapter 5).

8.2.2 Mass model

We combine these imaging and kinematic datasets to construct a model in which we are able to disentangle the dark and luminous contributions to the total mass profile of each EEL. For the main analysis, we consider two variations of one distinct mass model, as detailed below; however, in Section 8.4, we also explore alternative models in order to investigate the robustness and limitations of our inference.

We treat the mass density $\rho_{tot}(r)$ of the lens galaxy as the sum of a dark matter halo and a stellar component:

$$(8.1) \quad \rho_{tot}(r) = \rho_{DM}(r) + \rho_{\star}(r).$$

We do not include a black hole because realistic black hole masses are orders of magnitude smaller than the total Einstein mass and therefore have undetectable effects on the lensed features. Indeed, we investigate the effects of neglecting the black hole in Section 8.4.3 and show that it has a negligible impact on the inferred halo and stellar mass parameters.

To allow for astrophysical changes to the halo structure relative to the dark-matter-only NFW prediction, we model the halo as a modified gNFW profile:

$$(8.2) \quad \rho_{DM}(r) = \rho_0 \left(\frac{r}{r_s} \right)^{-\gamma} \left(1 + \frac{r^2}{r_s^2} \right)^{\frac{\gamma-3}{2}}$$

which is characterised by a scale radius r_s , inner slope γ and mass scale ρ_0 . The quadrature in the final bracket relative to the standard form for the gNFW profile (Equation 1.2) was introduced by Muñoz et al. (2001) to make the calculation of lensing deflection angles analytic, and only modifies the halo shape near the scale radius (where the profile transitions more sharply); our data only probe the mass at much smaller radii. To minimise degeneracies between ρ_0 , γ and r_s , we reparameterise the halo so that the mass scale is specified by $M_{DM}(R < 2.5\text{kpc})$, the *projected* dark matter mass within a circular 2.5 kpc aperture. We additionally assume that the halo is spherical and concentric with the stellar mass and light (though we also investigate the limitations of these assumptions in Section 8.4.3).

We parameterise the stellar mass with a stellar-mass-to-light ratio $Y_\star(R)$ such that the stellar surface mass density $\Sigma_\star(R)$ is related to the stellar surface brightness $I_\star(R)$ as

$$(8.3) \quad \Sigma_\star(R) = Y_\star(R) I_\star(R)$$

with $I_\star(R)$ described by either one or two Sérsic profiles and fixed by the inference in Chapter 5. The 3D mass density is then obtained by deprojecting $\Sigma_\star(R)$ assuming axisymmetry.

We consider two different models for $Y_\star(R)$. In our first model, we assume the stellar-mass-to-light ratio is spatially uniform:

$$(8.4) \quad Y_\star = \text{constant.}$$

The second model treats the stellar mass-to-light ratio as a power-law function of projected radius, such that

$$(8.5) \quad Y_\star = Y_{\star,1} R^\alpha$$

where $Y_{\star,1}$ is the stellar mass-to-light ratio at a radius of 1 kpc. We note that, although the limiting behaviour of this model is unrealistic, our data only probe the central regions of each galaxy, which makes its large-radius behaviour unimportant; equally, our implementation of this profile (by decomposing $Y_\star(R)$ into a sum of Gaussians; see Section 8.2.3) prevents $Y_\star(R \rightarrow 0)$ from becoming infinite. In the remainder of this chapter, we refer to these constant- and varying- stellar-mass-to-light ratio models as the CML and VML models respectively.

8.2.3 Lens modelling

We make full lensing reconstructions of the HST/ACS imaging using an extension of the methods presented in Chapter 5. Using the results of that study, which described the total

mass with a power-law profile and inferred the light distributions of the source and lens simultaneously, we subtract the flux contribution of the lens; we then parameterise the source light using either one or two Sérsic profiles, and calculate the deflection angles based on the mass models presented in Section 8.2.2 to form the source in the image plane. We also allow for an external shear to account for tidal perturbations, described by a magnitude Γ and direction θ_Γ . We use an unsaturated star in the HST field as a PSF, which we convolve with the model image to allow comparison with the data. For a given set of non-linear parameters describing the lens mass and source light profiles, we then determine the best combination of amplitudes for the source light and a uniform background component using a least-squares linear inversion. We model the V and I bands simultaneously (i.e. dictating that the luminous source structure is the same in both bands), allowing for a spatial offset between them due to image registration uncertainties, and calculate the contribution to the likelihood of the data \vec{D} given the model \vec{M} as

$$(8.6) \quad \ln L(\vec{D}|\vec{M}) = -\frac{1}{2} \sum_i \left(\frac{d_i - m_i}{\beta_f n_i} \right)^2 - \ln(\beta_f n_i)$$

where d_i , m_i and n_i are the i^{th} pixel in the data image, model image and noise map respectively, and β_f is a weighting term for the f^{th} filter which the model determines in order to maximise the combined posterior of the two lensing datasets and the dynamical data. The sum is over all unmasked pixels (for some systems, bright interloping objects are masked by hand).

For the CML models, we precalculate the deflections for the stellar mass using the Sérsic surface brightness profiles of Chapter 5; these can then be simply rescaled by Y_\star during the inference. For the VML models, where the stellar surface mass density is the product of a Sérsic profile and a power law, we make the deflection angle computations tractable by fitting the Sérsic profiles with multi-Gaussian expansions (MGEs) and precalculating the deflections for each MGE component individually; these can then be scaled by the value of $Y_\star(R)$ at the width σ_k of the k^{th} Gaussian; in this setup, the finite width of the innermost Gaussian prevents $Y_\star(R \rightarrow 0)$ from becoming infinitely large. We verify that our MGE decomposition gives an accurate description of the ‘true’ surface mass density for each lens.

8.2.4 Dynamical modelling

For a given mass model, we use the spherical Jeans equation (see e.g. Mamon & Łokas, 2005) to calculate the stellar velocity dispersion within a circular aperture of radius 1 arcsec assuming isotropic orbits. We investigate the possibility of allowing the anisotropy to take a

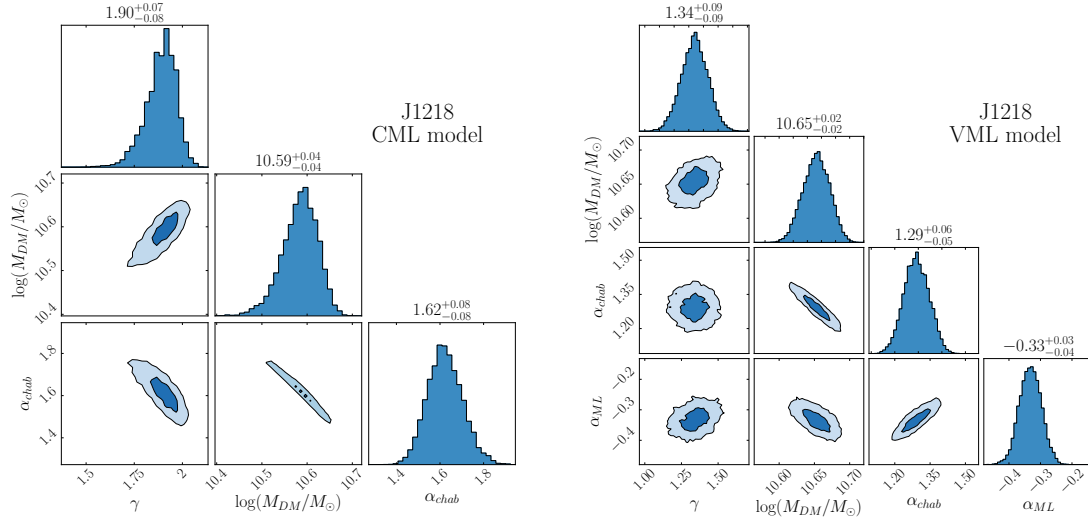


Figure 8.1: Summary of our inference on the mass model parameters for a typical EEL (J1218). For the CML model (left), we show our 2D and 1D marginalised inferences on the inner halo slope γ , the projected dark matter mass within 2.5 kpc, $\log(M_{DM}/M_\odot)$, and the IMF mismatch parameter α_{chab} . For the VML model (right), we additionally include our inference on the stellar-mass-to-light ratio slope α_{ML} .

non-zero but spatially constant value, but find that our data are not sufficient to constrain this extra parameter. We also note that the spherical assumption is an approximation given that the EELs lenses have light distributions that are described by elliptical Sérsic profiles; however, Sonnenfeld et al. (2015) showed that approximating a Sérsic profile with $q = 0.85$ using the spherical Jeans equation in this way only gives rise to an uncertainty of a few kms^{-1} .

The likelihood of the observed velocity dispersion given the model is then measured in a chi-squared sense as

$$(8.7) \quad \ln(\vec{D}|\vec{M}) = -\frac{1}{2} \left(\frac{\sigma_d - \sigma_m}{\sigma_n} \right)^2$$

for observed velocity dispersion σ_d , model velocity dispersion σ_m and uncertainty σ_n . Note that the use of the uncertainty weighting terms β_f in the lens modelling prevents the latter from overwhelming the likelihood calculation.

We combine dynamical and lensing likelihood terms to explore the posterior probability distribution of the model given the data using MCMC sampling, as implemented in EMCEE.

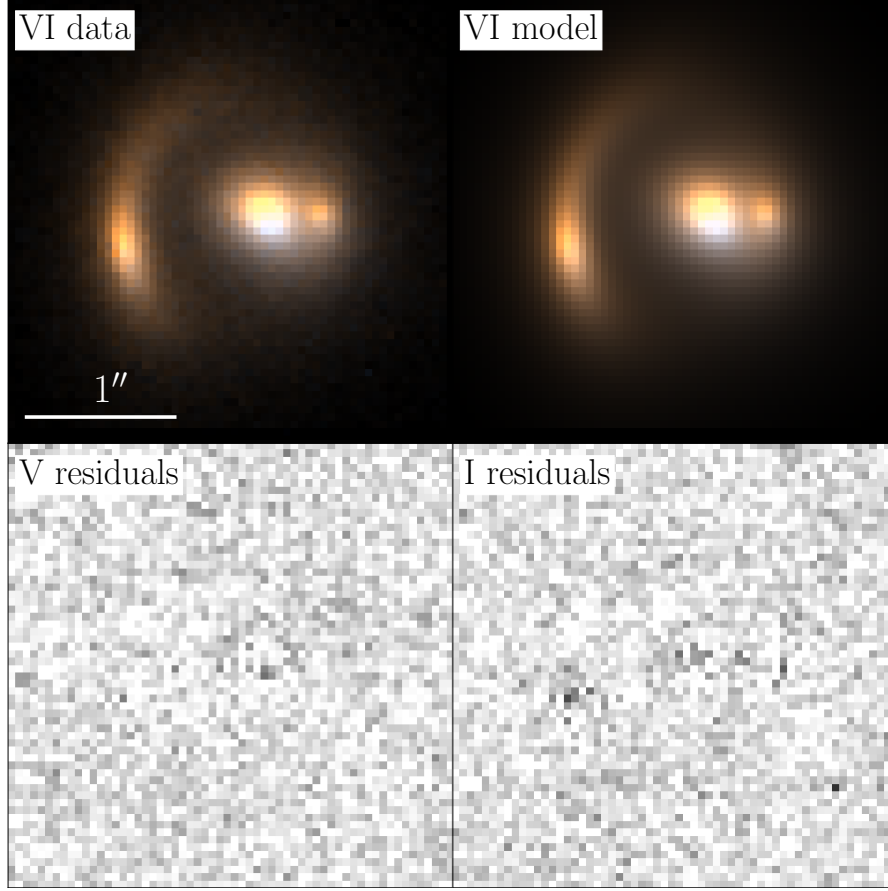


Figure 8.2: Reconstructed image for a typical EEL (J1218). We use the V and I HST images (top left) to make inference on the dark and luminous mass structure of the lens; our best model (top right) describes the data virtually down to the noise (bottom).

8.3 Results

Figure 8.1 shows our inference on the mass parameters for a typical lens (J1218) for both the CML and VML models, and Figure 8.2 shows the image, model and residuals for the VML model (the CML residuals are indistinguishable by eye). Tables 1 (CML) and 2 (VML) present our inferences on the mass structure for the full lens sample. In the following sections, we present these results in detail.

8.3.1 Models with a spatially constant Υ_\star

Table 8.1 summarises our inference on the mass parameters of the CML model, and Figure 8.1 (left) shows the posterior probability distributions for the parameters of interest

for a typical lens.

8.3.1.1 Stellar mass

The stellar-mass-to-light ratio Y_\star inferred from lensing and dynamics gives a stellar mass $M_{\star,LD} = Y_\star I_\star(R)$, which we use to test for deviations of the stellar populations from a Milky-Way-like IMF. We use the stellar masses $M_{\star,SPS}$ presented in Chapter 5, which were calculated via stellar population modelling of the photometry under the assumption of a Chabrier IMF (but with metallicity, age, reddening and the time constant of an exponentially decaying star formation history as free parameters), and calculate the IMF mismatch parameter

$$(8.8) \quad \alpha_{chab} = \frac{M_{\star,LD}}{M_{\star,SPS}}.$$

(Note that, with this definition, stellar populations with Chabrier and Salpeter IMFs have $\alpha_{chab} = 1$ and $\alpha_{chab} \sim 1.7$ respectively). The mismatch parameters for the CML models are included in Table 8.1 and Figures 8.1 (left) and 8.3 (top left). We characterise the sample by modelling our inferences on α_{chab} for individual EELs as being drawn from a Gaussian distribution $\alpha_{chab} \sim N(\mu_{ac}, \tau_{ac}^2)$. This constitutes a hierarchical model with hyperparameters $\omega = (\mu_{ac}, \tau_{ac})$ whose posterior probability distribution, given the data D , is given by

$$(8.9) \quad \begin{aligned} P(\omega|D) &\propto P(\omega)P(D|\omega) \\ &= P(\omega) \prod_i \int d\alpha_c P_i(D_i|\alpha_c) N(\alpha_c; \omega) \\ &\propto P(\omega) \prod_i \int d\alpha_c \frac{P_i(\alpha_c|D_i)}{P_i(\alpha_c)} N(\alpha_c; \omega) \end{aligned}$$

where we have shortened $\alpha_{chab} = \alpha_c$ for clarity. The first term $P(\omega)$ is the prior probability on ω , which we assume to be uniform; $P_i(\alpha_c)$ is the prior on α_{chab} for the i^{th} EEL, which is uniform and the same for all EELs; $P_i(\alpha_c|D_i)$ is the posterior on α_{chab} for the i^{th} EEL as inferred from our lensing and dynamical modelling. Equation 8.9 is therefore a product of i integrals of the trial parent distribution multiplied by the posterior on α_{chab} for the i^{th} lens, which we perform via Monte Carlo integration and sample with the same MCMC methods as in previous sections to find $\mu_{ac} = 1.69^{+0.17}_{-0.18}$, $\tau_{ac} = 0.62^{+0.20}_{-0.13}$ (Table 8.3); moreover, $\mu_{ac} > 1$ with 99% confidence. The sample of EELs lenses therefore strongly prefers Salpeter-like IMFs over Chabrier-like ones, consistently with previous results from lensing and dynamics (e.g. Auger et al., 2010b; Cappellari et al., 2012).

Finally, we model a linear relation between the stellar velocity dispersion σ_* and the mismatch parameter as $\alpha_{chab} = \beta(\sigma_*/200\text{kms}^{-1}) + \delta$ with an intrinsic dispersion v and an underlying parent distribution for σ_* given by $\sigma_* \sim N(\mu, \tau^2)$ (see Kelly, 2007) but find that, though our inference prefers $\beta > 0$, it is consistent with no correlation within 1σ (Table 8.4 and Figure 8.4). In a future work, we will extend our analysis to a larger sample of lenses, spanning a large range in σ_* , in order to place stronger constraints on this relation. We note, however, that our sample exhibits a large scatter in α_{chab} in spite of its small range in σ_* . Whilst some previous lensing studies assuming more rigid forms for the dark matter structure (i.e. NFW haloes; Treu et al., 2010) have found a positive correlation between α_{chab} and σ_* for other ETG samples, Auger et al. (2010b) has shown that this trend is removed when haloes are allowed to be adiabatically contracted. Regardless of whether a *correlation* exists, it is possible that the large α_{chab} *scatter* in our sample is also due to the increased flexibility of our halo models.

8.3.1.2 Dark matter

For two lenses (J0913 and J1606), the halo makes a negligible contribution to the Einstein mass, such that the halo properties are virtually unconstrained (Figure 8.3, upper right). The majority of the remaining lenses have extremely cuspy haloes. To make a meaningful comparison between systems with different scale radii and inner slopes, we calculate the mass-weighted slope within the effective radius as

$$(8.10) \quad \begin{aligned} \gamma' = \gamma(R < R_e) &= -\frac{1}{M_{DM}(r)} \int_0^{R_e} 4\pi r^2 \rho_{DM}(r) \frac{d \log \rho_{DM}}{d \log r} dr \\ &= 3 - \frac{4\pi R_e^3 \rho_{DM}(R_e)}{M(R_e)} \end{aligned}$$

for dark halo mass $M_{DM}(r)$ and density $\rho_{DM}(r)$ (e.g. Newman et al., 2015). We then model the EELs population as having mass-weighted inner slopes drawn from a Gaussian distribution $\gamma' \sim N(\mu_{\gamma'}, \tau_{\gamma'}^2)$, as in Section 8.3.1.1, and find $\mu_{\gamma'} = 2.25 \pm 0.20$, $\tau_{\gamma'} = 0.38^{+0.23}_{-0.12}$ (Table 8.3). We use the adiabatic contraction model of Gnedin et al. (2004) to calculate the halo inner slope that would result from a scenario in which both components begin as NFW profiles and the baryonic mass evolves into the inferred profile, and find that the typical halo inner slope that we would expect is $\gamma \sim 1.7$. This is inconsistent with our inference on $\mu_{\gamma'}$ under the assumptions of the CML model at the 3σ level, and suggests that either our model is inappropriate, or some additional physical process is acting on these haloes to contract them to such a large degree. Indeed, this strongly motivates the VML models below.

We note, however, that two systems (J1125 and J1323) are clear outliers from this trend and appear to have cored haloes. In Chapter 5, we noted that J1323 is unusual amongst the EELs as the lens is the bulge of a galaxy that has clear spiral structure at large radii; it also has an anomalously small central stellar velocity dispersion. It is therefore possible that this is a fundamentally different type of galaxy from the other EELs lenses and so may be primarily subject to different physical processes in its central regions. J1125 is unusual in that, in addition to a dark halo core, it also has a high stellar mass, implying an IMF heavier than Salpeter. It is possible that the high stellar mass of this system may also be responsible for its different halo structure – for instance, the high stellar mass may result from a large number of accretion events, during which dynamical heating may have reversed any initial halo contraction. It is difficult to draw conclusions about the population from a sample of 12 systems, but future work extending these methods to larger samples will allow a better understanding of the variation in halo structures that the EELs lenses exhibit.

8.3.2 Models with a spatially varying Υ_\star

A possible explanation for the cuspieness of the dark matter haloes that we infer with the CML models may be the presence of negative radial gradients in the stellar-mass-to-light ratios, due to trends in stellar population properties such as the age, metallicity and IMF (see Chapter 7). As the CML models do not allow for such a gradient, they may be forcing the halo profile to be steep in order to reproduce the slope of the *total* mass profile. We investigate this idea with our VML models, using a uniform prior on the stellar-mass-to-light ratio gradient α_{ML} over the range $(-1, 1)$ so as to be agnostic about its form.

8.3.2.1 Stellar mass

The VML and CML models are nested models, such that the VML model becomes a CML model when $\alpha_{ML} = 0$. However, as shown in Table 8.2 and Figure 8.3 (bottom), only two (four) out of our twelve systems have posterior distributions in which α_{ML} is consistent with zero at the 1σ (3σ) level. Two systems (J1125 and J1347) have mild positive gradients, and the remaining systems all have negative stellar-mass-to-light ratio gradients at $> 4\sigma$. Similarly to Section 8.3.1.1, we construct a hierarchical model for the $\alpha_{ML} \sim N(\mu_{am}, \tau_{am}^2)$ distribution across the sample, and find $\mu_{am} = -0.22 \pm 0.08$ and $\tau_{am} = 0.28^{+0.09}_{-0.06}$, confirming this initial impression. Whilst our current data do not allow us to disentangle the effects of age, metallicity and IMF, we note that the existence of negative stellar-mass-to-light ratio

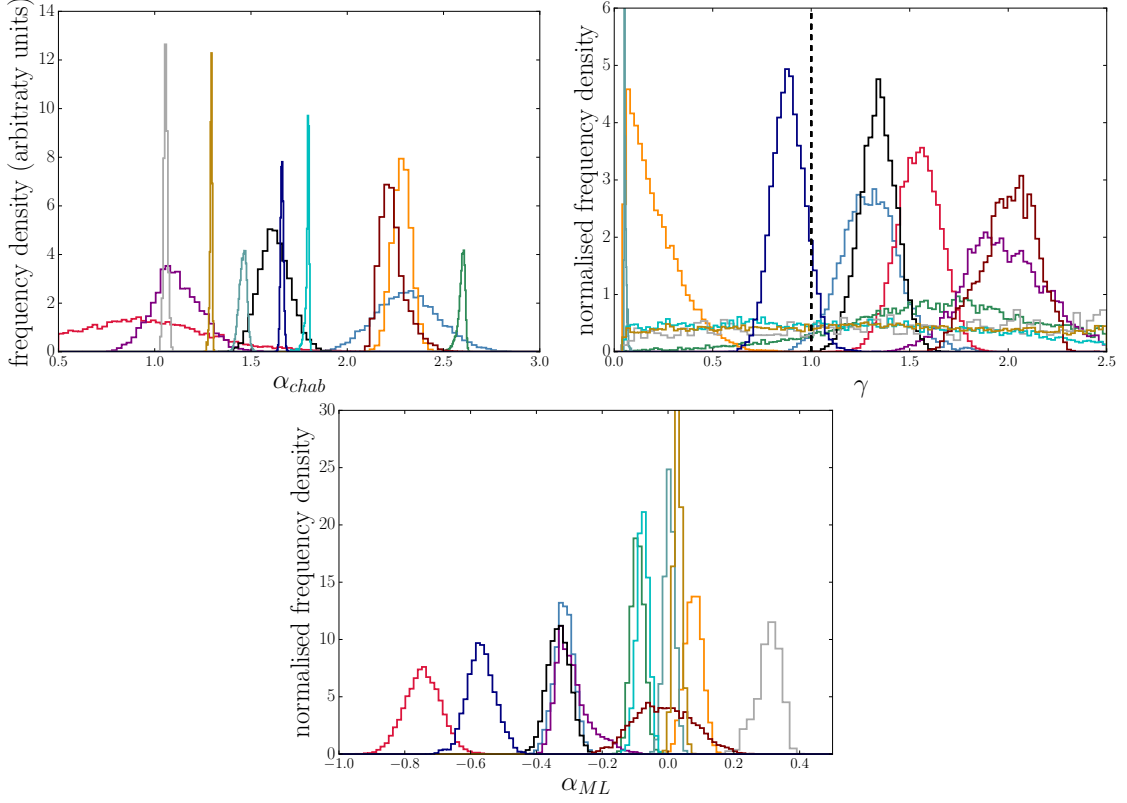


Figure 8.3: Marginalised inference on the IMF mismatch parameter α_{chab} (CML; upper left), halo inner slope γ (VML; upper right) and stellar-mass-to-light ratio slope α_{ML} (VML; bottom) for the EELs lens sample. Note that the α_{chab} histograms are arbitrarily normalised to reduce the dynamic range of the y-axis and so allow a straightforward visual comparison between different systems; the histograms in the other two panels are normalised to unity. The population is characterised by high stellar masses – implying stellar populations with IMFs which are heavier than Chabrier – steep haloes, and negative stellar-mass-to-light ratio gradients. These structural properties may suggest an evolutionary scenario in which adiabatic contraction of the halo and the inside-out growth of the stellar material are dominant.

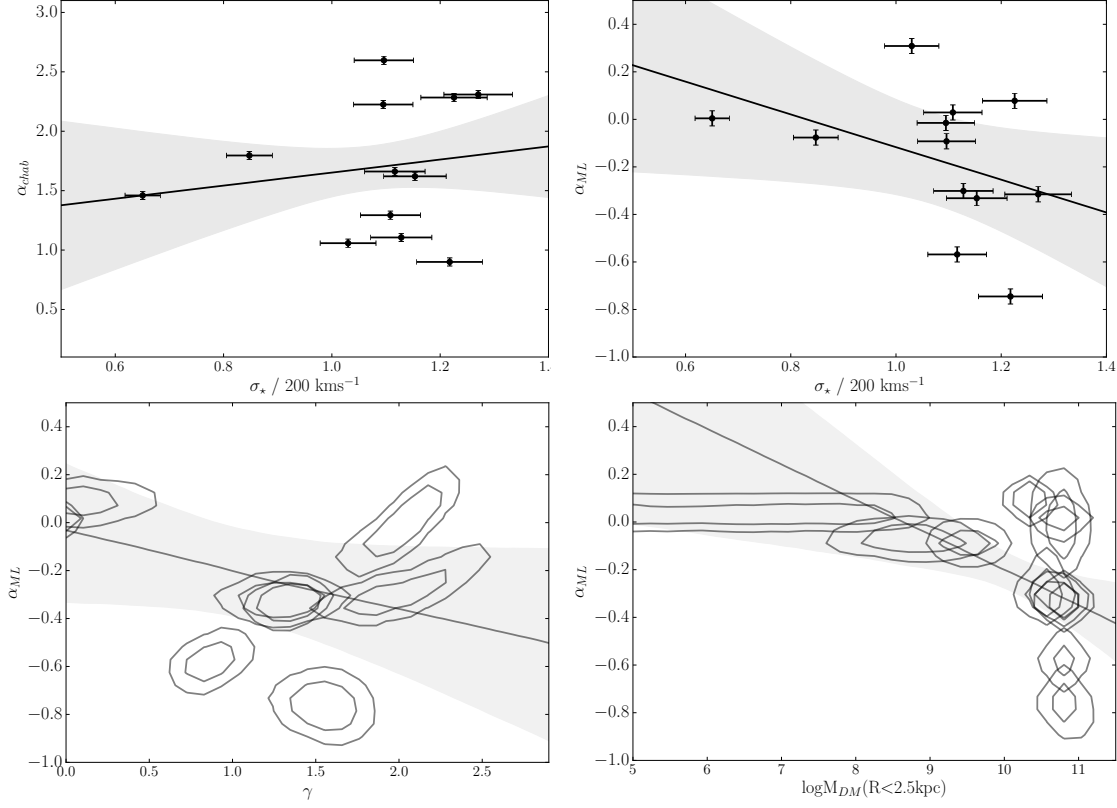


Figure 8.4: Hints at correlations between stellar mass and dark halo properties. We cannot constrain the correlation between stellar-mass-to-light ratio α_{chab} and stellar velocity dispersion σ_* (upper left) with any significance due to the small σ_* range spanned by our sample. However, we constrain linear relations between the stellar-mass-to-light ratio slope α_{ML} and the stellar velocity dispersion σ_* (upper right), halo inner slope γ (lower left) and projected dark matter mass within 2.5kpc, $\log M_{DM}(R < 2.5 \text{ kpc})$, (lower right) to have negative slopes at the 1σ level. This may be tentative evidence that the stellar population properties vary more strongly across galaxies residing in deeper gravitational potential wells. However, our current sample is too small and has insufficient dynamic range to allow any strong conclusions from these results. (Note that systems for which γ is virtually unconstrained are removed from the panel showing $\alpha_{ML} - \gamma$ for clarity, but were included in the inference.)

gradients is consistent with a scenario in which the IMF is bottom-heavy in the central regions and becomes increasingly Milky-Way-like at large radii, in line with recent results from some stellar population modelling studies (Martín-Navarro et al., 2015a; van Dokkum et al., 2016, and Chapter 7).

Though the dynamic range of our sample is small, we are able to make inference on a linear relation between the stellar velocity dispersion σ_* and the stellar-mass-to-light ratio slope α_{ML} – which, following Section 8.3.1.1, we model as $\alpha_{ML} = \beta(\sigma_*/200\text{kms}^{-1}) + \delta$ with an intrinsic dispersion ν and an underlying parent distribution for σ_* given by $\sigma_* \sim N(\mu, \tau^2)$. As shown in Figure 8.4 (left) and Table 8.4, we find that these properties are anticorrelated (at the 1σ level) such that the steepness of the stellar-mass-to-light ratio increases with velocity dispersion. This is consistent with the study of Martín-Navarro et al. (2015a), in which the IMF slope was found to vary radially in the two high-mass ($\sigma_* \sim 200\text{kms}^{-1}$) ETGs in their sample but not in the low-mass ($\sigma_* \sim 100\text{kms}^{-1}$) system. Since the velocity dispersion is a proxy for central galaxy mass, this is tentative evidence for an association between the strength of the stellar-mass-to-light ratio gradient and the depth of the central gravitational potential well. However, as the σ_* range of our sample is limited, we are not able to make strong statements about this hypothesis, and measurements of α_{ML} across ETGs which more representatively sample the σ_* space are now needed.

8.3.2.2 Dark matter

Our inference on the halo structure for the VML models is summarised in Table 8.2 and Figures 8.1 (right) and 8.3 (upper right). Compared with the CML model, VML haloes are universally less cuspy, and consistent with being drawn from a Gaussian population with mean mass-weighted slope within the effective radius $\mu_{\gamma'} = 1.66 \pm 0.11$ and dispersion $\tau_{\gamma'} = 0.26^{+0.13}_{-0.08}$ (Table 8.4). As in Section 8.3.1.2, we calculate our expectations for adiabatically contracted haloes, given the inferred baryonic mass distributions, and find that the negative stellar-mass-to-light ratio gradients (i.e. the cuspiest *stellar* mass profiles) lead to more contracted haloes $\gamma \sim 1.85$, which is consistent with our inference on $\mu_{\gamma'}$ within 2σ . The fact that the dark and luminous mass structure can now be explained consistently in the context of a simple physical model is not evidence that the mass models we are using are correct; however, together with the nested nature of the CML and VML models, it does suggest that the extra flexibility in the latter is meaningful.

EEL	$\log(M_{DM}/M_\odot)$	r_s (kpc)	γ	$M_\star (\times 10^{11} M_\odot)$	α_{chab}	σ_\star (kms $^{-1}$)
J0837	$10.72^{+0.04}_{-0.05}$	$96.17^{+84.27}_{-66.78}$	$2.27^{+0.03}_{-0.03}$	$2.85^{+0.20}_{-0.20}$	$2.31^{+0.17}_{-0.16}$	254.1 ± 12.7
J0901	$10.90^{+0.05}_{-0.05}$	$13.42^{+49.14}_{-3.57}$	$2.31^{+0.03}_{-0.03}$	$0.55^{+0.22}_{-0.23}$	$0.90^{+0.35}_{-0.37}$	243.5 ± 12.2
J0913	$8.92^{+0.33}_{-0.34}$	$112.92^{+70.54}_{-77.40}$	$1.38^{+0.46}_{-0.79}$	$2.96^{+0.02}_{-0.02}$	$2.60^{+0.01}_{-0.02}$	219.2 ± 11.0
J1125	$10.42^{+0.03}_{-0.02}$	$191.99^{+26.06}_{-41.76}$	$0.21^{+0.17}_{-0.11}$	$6.23^{+0.13}_{-0.14}$	$2.28^{+0.05}_{-0.05}$	245.1 ± 12.3
J1144	$10.74^{+0.04}_{-0.06}$	$71.15^{+89.73}_{-57.91}$	$2.34^{+0.05}_{-0.04}$	$1.30^{+0.16}_{-0.13}$	$1.11^{+0.14}_{-0.11}$	225.6 ± 11.3
J1218	$10.59^{+0.04}_{-0.04}$	$6.91^{+2.44}_{-1.20}$	$1.90^{+0.07}_{-0.08}$	$2.37^{+0.12}_{-0.11}$	$1.62^{+0.08}_{-0.08}$	230.7 ± 11.5
J1323	$10.60^{+0.01}_{-0.01}$	$184.41^{+1.21}_{-2.40}$	$0.05^{+0.01}_{-0.00}$	$1.83^{+0.02}_{-0.03}$	$1.46^{+0.02}_{-0.02}$	130.1 ± 6.5
J1347	$10.02^{+0.03}_{-0.03}$	$1.56^{+0.16}_{-0.10}$	$2.78^{+0.01}_{-0.02}$	$1.75^{+0.02}_{-0.02}$	$1.06^{+0.01}_{-0.01}$	206.0 ± 10.3
J1446	$8.04^{+0.71}_{-1.04}$	$94.08^{+59.93}_{-61.84}$	$1.41^{+0.80}_{-0.92}$	$1.42^{+0.00}_{-0.01}$	$1.80^{+0.01}_{-0.01}$	169.5 ± 8.5
J1605	$10.19^{+0.05}_{-0.09}$	$24.94^{+103.17}_{-14.25}$	$1.99^{+0.10}_{-0.13}$	$2.57^{+0.09}_{-0.06}$	$2.22^{+0.08}_{-0.05}$	219.0 ± 11.0
J1606	$7.00^{+1.58}_{-1.33}$	$120.82^{+61.80}_{-71.25}$	$1.67^{+0.73}_{-0.98}$	$2.83^{+0.01}_{-0.01}$	$1.29^{+0.00}_{-0.01}$	221.6 ± 11.1
J2228	$9.65^{+0.03}_{-0.03}$	$1.62^{+0.17}_{-0.14}$	$2.76^{+0.02}_{-0.03}$	$2.87^{+0.01}_{-0.01}$	$1.66^{+0.01}_{-0.01}$	223.3 ± 11.2

Table 8.1: CML models for all the EELs. From left to right, the columns show the system name, the projected dark matter mass within a 2.5kpc aperture, the halo scale radius, the halo inner slope, total stellar mass, IMF mismatch parameter and stellar velocity dispersion measured in an aperture of 1 arcsec radius; uncertainties are statistical only. In the context of this model, most systems have high stellar masses and steep halo slopes (see Table 8.3).

EEL	$\log(M_{DM}/M_\odot)$	r_s (kpc)	γ	$M_\star (\times 10^{11} M_\odot)$	α_{ML}	α_{chab}
J0837	$10.62^{+0.04}_{-0.03}$	$13.25^{+3.81}_{-1.58}$	$1.30^{+0.14}_{-0.14}$	$3.04^{+0.14}_{-0.16}$	$-0.32^{+0.03}_{-0.03}$	$2.46^{+0.12}_{-0.13}$
J0901	$10.78^{+0.03}_{-0.03}$	$9.70^{+1.13}_{-0.43}$	$1.55^{+0.11}_{-0.11}$	$0.76^{+0.10}_{-0.10}$	$-0.75^{+0.05}_{-0.05}$	$1.24^{+0.17}_{-0.17}$
J0913	$9.47^{+0.20}_{-0.18}$	$120.01^{+65.90}_{-77.91}$	$1.66^{+0.46}_{-0.49}$	$2.85^{+0.03}_{-0.04}$	$-0.09^{+0.02}_{-0.02}$	$2.50^{+0.03}_{-0.04}$
J1125	$10.29^{+0.06}_{-0.06}$	$185.40^{+31.49}_{-47.19}$	$0.18^{+0.16}_{-0.09}$	$6.32^{+0.13}_{-0.13}$	$0.08^{+0.03}_{-0.03}$	$2.32^{+0.05}_{-0.05}$
J1144	$10.45^{+0.05}_{-0.06}$	$25.90^{+117.63}_{-13.76}$	$1.97^{+0.23}_{-0.18}$	$1.85^{+0.09}_{-0.08}$	$-0.30^{+0.06}_{-0.03}$	$1.57^{+0.08}_{-0.07}$
J1218	$10.65^{+0.02}_{-0.02}$	$9.51^{+0.46}_{-0.18}$	$1.34^{+0.09}_{-0.09}$	$1.89^{+0.08}_{-0.08}$	$-0.33^{+0.03}_{-0.04}$	$1.29^{+0.06}_{-0.05}$
J1323	$10.61^{+0.01}_{-0.01}$	$184.52^{+1.10}_{-2.61}$	$0.05^{+0.01}_{-0.00}$	$1.80^{+0.03}_{-0.03}$	$0.00^{+0.02}_{-0.02}$	$1.44^{+0.03}_{-0.02}$
J1347	$6.74^{+1.30}_{-1.15}$	$114.55^{+64.17}_{-66.15}$	$1.38^{+0.78}_{-0.86}$	$1.79^{+0.03}_{-0.03}$	$0.31^{+0.03}_{-0.04}$	$1.09^{+0.02}_{-0.02}$
J1446	$8.74^{+0.41}_{-0.46}$	$98.56^{+57.87}_{-60.26}$	$1.11^{+0.78}_{-0.71}$	$1.38^{+0.01}_{-0.02}$	$-0.08^{+0.02}_{-0.02}$	$1.74^{+0.01}_{-0.02}$
J1605	$10.65^{+0.04}_{-0.06}$	$14.45^{+39.25}_{-4.52}$	$2.01^{+0.13}_{-0.15}$	$1.68^{+0.16}_{-0.13}$	$-0.01^{+0.09}_{-0.09}$	$1.46^{+0.14}_{-0.11}$
J1606	$6.85^{+1.19}_{-1.26}$	$102.44^{+70.34}_{-62.54}$	$1.26^{+0.81}_{-0.80}$	$2.59^{+0.01}_{-0.01}$	$0.03^{+0.01}_{-0.01}$	$1.18^{+0.00}_{-0.00}$
J2228	$10.70^{+0.02}_{-0.03}$	$7.84^{+0.46}_{-0.20}$	$0.88^{+0.09}_{-0.08}$	$1.24^{+0.12}_{-0.10}$	$-0.57^{+0.04}_{-0.04}$	$0.71^{+0.07}_{-0.06}$

Table 8.2: VML models for all the EELs. From left to right, the columns show the system name, the projected dark matter mass within a 2.5kpc aperture, the halo scale radius, the halo inner slope, total stellar mass, stellar-mass-to-light ratio slope and IMF mismatch parameter; uncertainties are statistical only. In the context of this model, most systems have high stellar masses, steep halo slopes and stellar-mass-to-light ratio gradients which are negative or consistent with zero (see Table 8.3).

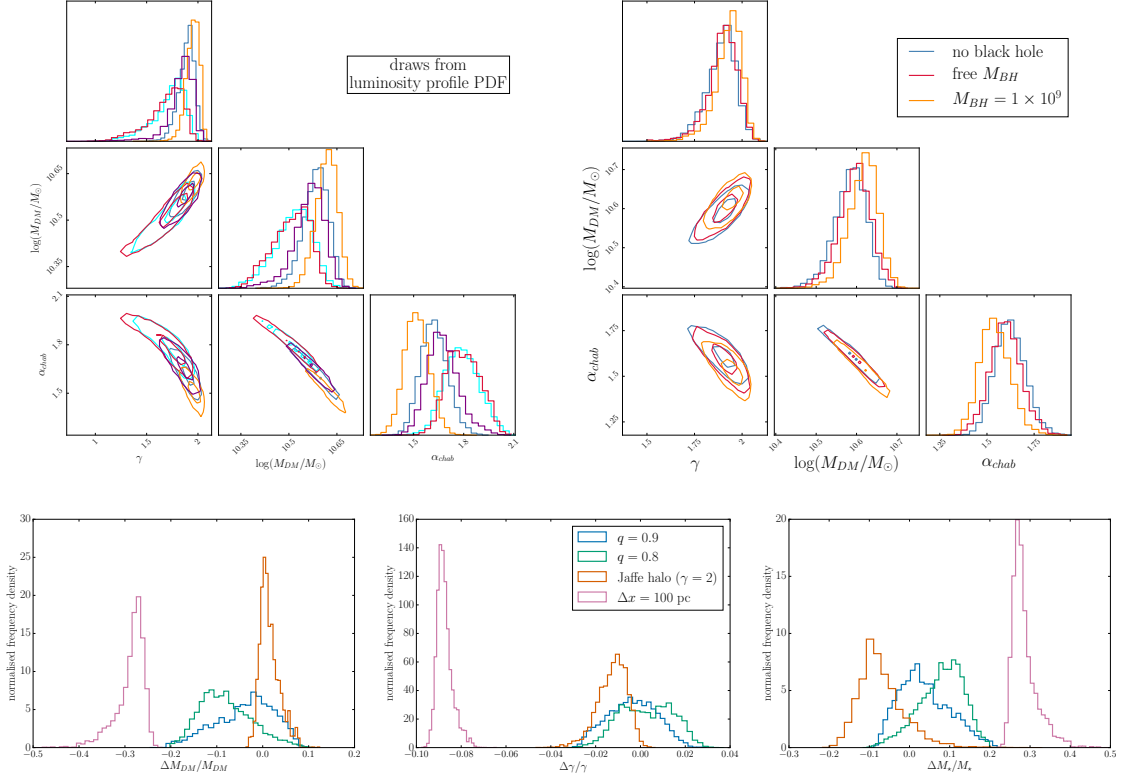


Figure 8.5: Systematics in the construction of lens models for the EELs. Drawing from the posterior distribution of the lens light parameters introduces additional uncertainty on the mass model parameters that are of the order to the statistical uncertainties (upper left); neglecting a massive central black hole has negligible impact on the inference (upper right). Assuming the halo is spherical introduces a small systematic bias into the inference, whereas assuming the halo is concentric with the stellar mass could introduce a larger bias (lower row).

model	X	μ_X	τ_X
CML	α_{chab}	$1.69^{+0.17}_{-0.18}$	$0.62^{+0.20}_{-0.13}$
CML	$\gamma(R < R_e)$	2.25 ± 0.20	$0.38^{+0.23}_{-0.12}$
VML	α_{ML}	-0.22 ± 0.08	$0.28^{+0.09}_{-0.06}$
VML	$\gamma(R < R_e)$	1.66 ± 0.11	$0.26^{+0.13}_{-0.08}$

Table 8.3: We characterise the EELs lens sample using a hierarchical model in which the property X of each EEL is drawn from a Gaussian parent distribution $X \sim N(\mu, \tau^2)$. In general, the EELs lenses have stellar masses consistent with Salpeter IMFs, dark halo slopes consistent with simple adiabatic contraction models, and negative stellar-mass-to-light ratio slopes.

relation	δ	β	ν	μ	τ
$\sigma_\star - \alpha_{chab}$	$0.57^{+1.19}_{-1.20}$	$1.08^{+1.30}_{-1.30}$	$0.63^{+0.19}_{-0.13}$	$1.07^{+0.06}_{-0.06}$	$0.18^{+0.05}_{-0.04}$
$\sigma_\star - \alpha_{ML}$	$-0.69^{+0.59}_{-0.58}$	$0.57^{+0.64}_{-0.64}$	$0.31^{+0.10}_{-0.06}$	$1.07^{+0.06}_{-0.06}$	$0.18^{+0.05}_{-0.04}$
$\gamma - \alpha_{ML}$	$-0.03^{+0.18}_{-0.20}$	$-0.16^{+0.15}_{-0.16}$	$0.29^{+0.09}_{-0.06}$	$1.16^{+0.22}_{-0.23}$	$0.72^{+0.23}_{-0.16}$
$\log M_{DM} - \alpha_{ML}$	$1.29^{+1.02}_{-1.08}$	$-0.15^{+0.11}_{-0.10}$	$0.26^{+0.09}_{-0.05}$	$10.11^{+0.29}_{-0.30}$	$0.96^{+0.34}_{-0.21}$

Table 8.4: For mass properties X and Y , we model a linear relation $\alpha_X = \beta Y + \delta$ with intrinsic dispersion ν and X drawn from a parent distribution $X \sim N(\mu, \tau^2)$; note that, for the velocity dispersion, we use $X = \sigma_\star/200\text{kms}^{-1}$. We find weak, tentative evidence that the stellar-mass-to-light ratio slope anticorrelates with other mass diagnostics such as the stellar velocity dispersion, halo inner slope and projected central halo mass. However, the small size and dynamic range of our sample means that we can only make inference at the 1σ level, and larger, more diverse samples will be necessary to test the robustness and significance of these trends.

8.3.3 The population as a whole

This study represents the first step towards the construction of precise, physically motivated mass models for a larger sample of lenses, with which it will be possible to probe correlations between different mass structure properties and thus the underlying evolutionary processes. With a sample of just 12 lenses which span a small range of luminous properties (for instance, their stellar masses span a range of only ~ 2 ; see their $R_e - M_\star$ relation in Chapter 5), the statements we can make about massive ETGs as a population are limited. Nevertheless, we make some early investigations here.

We have already presented models for the linear relations between the stellar velocity dispersion σ_\star and the stellar mass properties in Sections 8.3.1.1 and 8.3.2.1, and shown that the velocity dispersion may be anticorrelated with the stellar-mass-to-light ratio slope (Figure 8.4, top row). Here, we additionally consider relations between the stellar-mass-to-light ratio slope and the halo properties that were inferred in our lensing+dynamics

modelling framework. As in Section 8.3, we use the linear regression model of Kelly (2007) to fit the relation $Y = \beta X + \delta$ between mass properties X and Y , allowing an intrinsic dispersion v in the direction of Y and an underlying parent distribution $X \sim N(\mu, \tau^2)$. We use the MCMC samples from our inferences on individual lenses to treat the uncertainties and covariances precisely.

As shown in Figure 8.4 (bottom row) and Table 8.4, we find evidence (at the 1σ level) that α_{ML} becomes steeper with increasing inner halo slope γ and increasing central projected dark matter mass $\log M_{DM}(R < 2.5\text{kpc})$. As with the $\alpha_{ML} - \sigma_*$ relation, this suggests that the depth of the central potential well may play an important role in determining the baryonic mass structure. Larger samples are now needed in order to more conclusively establish whether these trends are significant.

8.3.4 Systematics

In the tables and figures in this chapter, we show the statistical uncertainties from our MCMC sampling; however, there are a number of systematic uncertainties introduced by our modelling assumptions. Indeed, this is evidenced by the difference between our inferences on the halo structure between the CML and VML models. In this section, we quantify some of these additional sources of uncertainty.

8.3.4.1 Stellar mass

Firstly, to make the lens modelling computationally feasible, we carry out our lensing and dynamical modelling with the stellar light distribution fixed according to the maximum-posterior profile that was inferred in Chapter 5. That study simultaneously modelled the light from the lens and the source such that the two components were robustly disentangled; nevertheless, the lens light distribution is only known as a posterior probability distribution with some finite width. By treating this distribution as a delta function, we are underestimating our uncertainty on the mass parameters. To quantify this extra uncertainty, we randomly draw lens light distributions from the posteriors of a subset of the lenses and rerun the inference (using the CML model for simplicity). Figure 8.5 (top left) shows the scatter in our inference on the mass parameters for 5 independent runs of the CML model for J1218 (note that we performed 15 runs overall, but show a representative sample here for clarity): we find additional uncertainties of ± 0.10 in α_{chab} , ± 0.09 in γ and ± 0.04 in $\log M_{proj}$. These are generally of the order of the statistical uncertainties on those parameters, and are typical across the sample.

8.3.4.2 Black holes

A second potential source of bias in our inference on the mass slopes of the dark and luminous components is the fact that we are ignoring the presence of the central black hole. Based on the $M_{BH} - \sigma_*$ relation (McConnell & Ma, 2013), the median velocity dispersion $\langle \sigma_* \rangle = 223 \text{ km s}^{-1}$ of the EELs lenses corresponds to a black hole mass $M_{BH} \sim 3.9 \times 10^8 M_\odot$. In contrast, the median Einstein mass of the EELs lenses is $\langle M_{Ein} \rangle = 1.1 \times 10^{11} M_\odot$; the black hole mass therefore makes a negligible ($\sim 0.3\%$) contribution to the total lensing mass.

Nevertheless, given that we infer steep total mass slopes, it is important to quantify any potential bias that could arise from ignoring this contribution. We therefore run two additional tests for a subset of lenses. First, we run a modified version of our CML models in which a black hole is included. We assume the black hole to be concentric with the halo and baryonic components and infer its mass along with the other mass parameters; in this case, we are unable to obtain meaningful constraints on the black hole mass and find it to be consistent with zero. Given the very small contribution of any realistic black hole to the projected mass within the Einstein radius, this is not a surprising result. Secondly, we rerun a set of CML models which now include black holes with fixed mass. To check extreme cases, we consider black hole masses $M_{BH} = (1, 5, 10, 50) \times 10^8 M_\odot$; in all these cases, our inference does not change within the statistical uncertainties. Figure 8.5 (top right) summarises these tests.

8.3.4.3 Flattened and offset haloes

A further possible source of bias is our assumption that the dark matter haloes are spherical and concentric with the light. It is possible that the halo may be flattened, though simulations suggest that the halo is generally rounder than the light (Abadi et al., 2010; Zemp et al., 2012). The median axis ratio for the EELs lens light profiles, $q = 0.8 \pm 0.1$, therefore sets a lower limit on the range of halo axis ratios that we might expect. The halo may also be offset from the stellar mass due to recent merger events: though CDM simulations constrain dark/light offsets to be less than the gravitational softening length of the simulations (350 pc; Schaller et al., 2015), offsets have been inferred to exist in some strong lensing galaxy clusters (Massey et al., 2015; Harvey et al., 2017).

Whilst we do not attempt to constrain the axis ratios or spatial offsets of the haloes in our sample – we are unable to make meaningful inference on the axis ratio, and allowing spatial offsets would prohibit the use of simple dynamical models – we construct a suite of simulated lenses to investigate the bias that is introduced by modelling flattened, offset

haloes using spherical, non-offset lens models. We generate synthetic high-signal-to-noise HST-like images of fake lens systems in which the source light is given by a single Sérsic profile and the lensing mass is the linear sum of (1) a stellar mass component with a spatially uniform stellar-mass-to-light ratio and a Sérsic light profile and (2) a gNFW halo which is flattened and/or offset from the stellar mass. We consider halo axis ratios in the range $q_h = (0.8, 1)$ to span the axis ratio range of the stellar light profiles of the EELs, and spatial offsets up to 0.1 kpc.

Figure 8.5 (bottom) shows the results of these tests. For the flattened, concentric case, we find that inference on M_\star , γ and $\log M_{DM}$ is robust against the simplifying assumption of a spherical halo. For offset haloes, on the other hand, we find that the accuracy of our inference on the mass parameters depends on the cuspieness of the halo but is relatively insensitive to the flattening. Our recovery of the halo slope deteriorates from $\Delta\gamma/\gamma = 0.04$ for cuspy haloes ($\gamma = 2$) to $\Delta\gamma/\gamma = 0.25$ for NFW haloes ($\gamma = 1$). Our recovery of $\log M_\star$ and $\log M_{DM}$ is worse than this, however. For cuspy haloes ($\gamma = 2$), we find $\Delta M_\star/M_\star \sim 0.3$ and $\Delta M_{DM}/M_{DM} \sim -0.3$, whereas for NFW haloes ($\gamma = 1$), we underestimate M_\star and overestimate M_{DM} by a similar fraction. If the haloes of the EELs lenses are significantly offset from the baryonic distributions, then, this could lead to a systematic bias in our results. We emphasise, however, that there is no theoretical or observational evidence for the existence of such offsets in field galaxies (Schaller et al., 2015), and that the BCGs in which offsets have been tentatively inferred occupy very different environments from the EELs lenses, which, as isolated systems, are likely to be more dynamically relaxed than BCGs at cluster centres.

8.3.4.4 Non-gNFW haloes

Finally, we use our simulations to test the bias introduced by modelling non-gNFW haloes using a gNFW profile. Whilst the NFW profile provides a good description of dark haloes in the absence of baryonic physics, simulations including prescriptions for baryonic processes which reproduce many of the observed scaling relations in galaxies on dwarf to Milky Way mass scales are better described by a yet more general form of the NFW profile in which the outer slope and break softening also vary (though the variation can be well parameterised by the stellar-to-halo mass ratio; Di Cintio et al., 2014). Comparable studies have not yet been carried out for massive ETGs (as it is much more computationally expensive to simulate such massive galaxies in large numbers), but it is unlikely that a gNFW profile can provide a complete description of the halo on these scales either. Furthermore, when we use gNFW profiles to fit the adiabatically contracted haloes that we calculated for the EELs

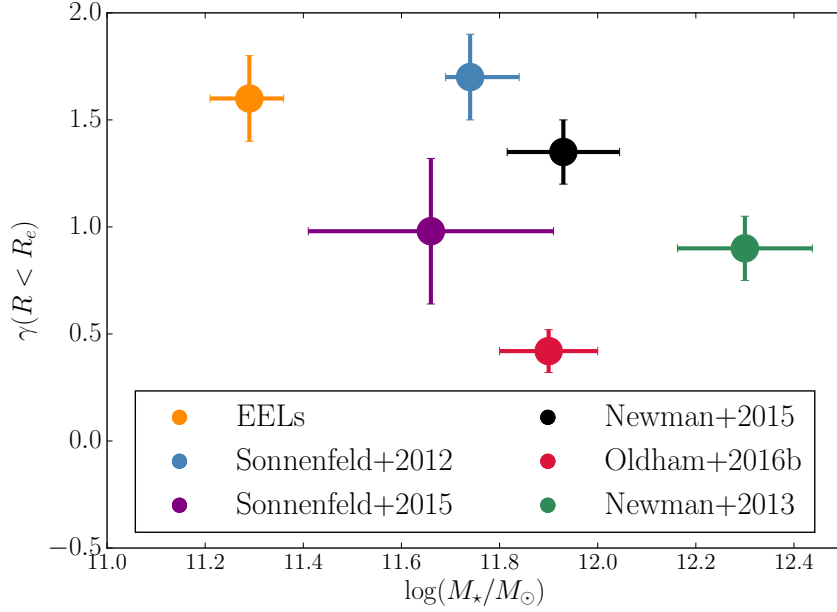


Figure 8.6: A summary of inferences on the stellar mass and dark matter halo slope for ETGs. The mean mass-weighted halo slope within the effective radius appears to become increasingly shallow in ETGs in denser environments. This may reflect a real environmental dependence of the relative importance of different baryonic processes, such that the dark halo is more significantly heated in dense environments where the rate of merging and accretion events is higher.

(Section 8.3), we find residuals of $\sim 20\%$ at intermediate radii; this is because adiabatic contraction acts to ‘pinch’ a pristine NFW halo such that the halo remains NFW-like in these regions, whereas the gNFW profile forces the halo slope to steepen monotonically with radius.

We test the robustness of our interpretation of γ as the inner slope of the dark matter profile by simulating lenses with haloes described by Jaffe profiles (with inner and outer slopes $\gamma = 2$ and $\gamma = 4$ respectively). We find our inference on γ using gNFW models recovers truth within 4% (Figure 8.5, bottom row). This indicates that γ is indeed recovering the inner halo slope, regardless of the slope at larger radii and the strength of the break. Though lensing is sensitive to the *projected* mass, the contribution of the outer regions, where the halo structure starts to differ substantially, is sufficiently small as to have a negligible effect on our inference on the properties of the central regions.

8.4 Discussion

We have presented models of the dark and luminous mass structure of 12 massive ETG lenses, and found evidence that the sample can be characterised as having contracted haloes and negative stellar-mass-to-light ratio gradients (which could be due to variations in the stellar age, metallicity or IMF). We now place our results in the context of previous work and consider their implications for the physics of ETG evolution.

8.4.1 Dark matter

We find a majority of the EELs lenses to have haloes that are centrally steeper than the NFW expectation but consistent with the predictions of simple models for adiabatic contraction.

8.4.1.1 Evidence for adiabatic contraction

Previous studies of the halo structure of massive isolated ETGs are few, but are nevertheless consistent with our result. Sonnenfeld et al. (2012) used the larger-radius constraints available for the rare double source plane lens J0946+1006 to constrain the halo slope of that system to be $\gamma = 1.7 \pm 0.2$, which they interpreted as evidence for adiabatic contraction due to the initial infall of gas; Grillo (2012) combined Einstein radius and velocity dispersion measurements with simple stellar population models for a sample of 39 strong lenses to constrain the average logarithmic density slope of the population, which they found to be steep but dependent on the assumed IMF ($\gamma = 2 \pm 0.2$ for a Chabrier IMF and 1.7 ± 0.5 for a Salpeter IMF). Additionally, Napolitano et al. (2010) found an anticorrelation between the central dark matter density and galaxy size which they interpreted as evidence for halo cuspieness. Our sample spans a range of halo inner slopes which is consistent with these previous conclusions, though we are able to constrain the inner slopes of individual galaxies with higher precision.

Furthermore, our inferences on the halo structure are quantitatively consistent with the predictions of the adiabatic contraction model of Gnedin et al. (2004), as discussed in Section 8.3. They are also qualitatively in agreement with the more realistic study of Duffy et al. (2010), which used high-resolution hydrodynamical simulations to measure the impact of adiabatic contraction on haloes in a cosmological context. That study found that the haloes of isolated galaxies become contracted such that $\gamma \sim 2$ in the presence of weak supernova feedback, with the action of either AGN or strong supernova feedback reducing this to $1.4 \leq \gamma \leq 2$.

8.4.1.2 Environment-dependent halo structure?

The fact that the majority of these massive isolated ETGs have contracted haloes motivates us to return to the suggestion, introduced by Newman et al. (2015), that halo structure may be an environment-dependent property; that is, that the relative importance of different baryonic processes may depend on environment.

Whilst studies of the dark haloes of BCGs have found these structures to have less mass centrally than the NFW prediction (Newman et al., 2013 and Chapters 3 and 7), the first (and, so far, only) study of group-scale haloes has found them to be mildly contracted (Newman et al., 2015). Finally, in this study we find the haloes of our sample of isolated ETGs to be strongly contracted, strengthening – and extending to lower masses, see Figure 8.6 – previous results on these scales (Sonnenfeld et al., 2012; Grillo, 2012). In Section 8.3.2.2, we inferred the mean mass-weighted halo slope within the effective radius, γ' , of the sample to be $\mu_{\gamma'} = 1.66 \pm 0.11$; we now combine this result with similar measures on cluster- and group-scales from Newman et al. (2013, 2015) and Chapter 3 and in the field from Sonnenfeld et al. (2012, 2015, though constraints from the latter are weak) in Figure 8.6 to show that the environmental dependence suggested by those studies at larger Einstein radii appears to persist on the scale of the isolated EELs (though with significant scatter).

This sequence of increasing halo expansion in increasingly dense environments suggests a scenario in which the dark matter haloes of ETGs become initially contracted due to the infall of gas, and are subsequently heated during the accretion events which also cause them to grow in size, with the degree to which this heating erases the initial contraction signature dependent on the amount of accretion that a particular galaxy experiences. Since ETGs in dense environments will typically accrete many more objects than those in the field, this leads to the trend we observe in which the inner halo slope becomes shallower in denser environments.

Based on current results, this scenario is compelling, but more studies on all mass scales are needed to test it further. We plan to extend the models presented here to a much larger sample of field ETG lenses in a future work.

8.4.1.3 Limitations at large radii

Finally, we note that the main limitation of our current modelling paradigm is that we are unable to recover the halo structure at large radii due to the absence of mass probes on these scales. The distribution of R_{Ein}/R_e for the EELs has a median of ~ 0.5 , such that the lensing only probes the mass in the central regions; though the velocity

dispersion is measured over a larger aperture ($\sim 1R_e$), this is luminosity-weighted which means that it is also most constraining in the centre. Since these are isolated systems, there are no substantial dynamical tracer populations such as satellite galaxies at larger radii, and the weak lensing signal or X-ray emission of an individual object at $z \sim 0.3$ would be prohibitively low. We are therefore unable to make meaningful inference on the halo scale radius r_s , and find that its posterior distribution either resembles the prior (which is uniform for $0 \leq r_s \leq 200\text{kpc}$; this is the result for most systems), or becomes unreasonably small (and comparable to the effective radius of the light). The former scenario is uninformative and prohibits inference on larger-radius mass measures such as the virial mass; the latter scenario leads to virial masses which are unreasonably low, and may be a sign of a mismatch between the true halo structure and the gNFW model that we are using to describe it (indeed, we see a similar effect in our simulations when the true and model haloes become particularly mismatched; see Section 8.3.4).

We note that previous strong lensing studies have also generally predicted halo masses below abundance matching expectations based on stellar population modelling of the galaxy colours (Auger et al., 2010b; Sonnenfeld et al., 2015). One option to make progress would be to use abundance matching expectations for the virial mass as a large-radius mass constraint; however, these relations have a large scatter and so do not have much constraining power. Moreover, the fact that we explicitly find stellar masses which require a heavier IMF than that assumed in the construction of abundance matching relations (i.e. heavier than a Chabrier IMF) makes the physical reasoning behind adding this constraint tenuous. We suggest that a better way to connect these high-precision measurements of the inner halo structure with constraints on larger scales will be statistically, by comparing inferences on samples of strong lenses such as the EELs with inferences on samples in which, for instance, the weak lensing signal is significant. It is likely that making this connection will also require the investigation and adoption of more flexible halo models than the gNFW profile.

8.4.2 Stellar mass and the IMF

All of the EELs lenses have stellar-mass-to-light ratios which are heavier than that expected assuming Milky-Way-like IMFs, and nearly all have negative stellar-mass-to-light ratio gradients. Though these gradients could be due to trends in the age, metallicity or IMF of the stellar populations, they are consistent with a picture in which the IMF is bottom-heavy in the central regions but Milky-Way-like in the outskirts as has been previously suggested (Martín-Navarro et al., 2015a; van Dokkum et al., 2016 and Chapter

7). We explore this idea further here.

8.4.2.1 Radially declining stellar-mass-to-light ratios

Over the past two years, a number of stellar population studies have attempted to place constraints on the existence of IMF gradients in ETGs, but no agreement has yet been reached. On the one hand, the initial study of three ETGs by Martín-Navarro et al. (2015a) found variations in the gravity-sensitive absorption features of the two massive ($\sigma_* \sim 200 \text{ km s}^{-1}$) ETGs in their sample that could not be explained unless the fraction of low-mass stars was allowed to decrease with increasing radius, and van Dokkum et al. (2016) modelled the full optical spectra of seven further ETGs to show that their IMFs also become more bottom-heavy centrally. However, a number of other studies – mostly using near-infrared absorption features – have found that radial trends in similarly massive ETGs can be explained by abundance variations, without the need for a changing IMF (Vaughan et al., 2016; Zieleniewski et al., 2017; Alton et al., 2017).

To help settle this question, independent constraints on ETG stellar-mass-to-light ratios from lensing and/or dynamical studies are needed. So far, however, only one other such study has been published, which used molecular gas kinematics to infer the total-mass-to-light ratios $\Upsilon_{tot}(R)$ across the central ~ 1 kpc of eight ETGs. This study found a large amount of scatter among the sample and no evidence for any systematic trends (though the systems with *rising* $\Upsilon_{tot}(R)$ profile were generally the least massive ones; Davis & McDermid, 2017). However, that study did not separate the dark and luminous mass structure, which may be responsible for some of the scatter that was found. Indeed, Chapter 7 of this thesis, which represents the first attempt to simultaneously model the dark matter halo and a stellar mass with a radially varying IMF, found strong evidence for the existence of a radially declining IMF slope in M87.

The study presented here reaches stronger conclusions about the existence of radially declining stellar-mass-to-light ratios for our sample of 12 strong lenses, and, thus, for systematic changes in the stellar population properties *across* the individual galaxies. So far, we have not been able to identify whether these gradients are being driven by IMF, metallicity or age variations, since our assumption in Chapter 5 that the lens galaxies do not exhibit colour gradients restricts our photometry to integrated colours. We can therefore only calculate stellar population masses $M_{*,SPS}$ assuming that all stellar population properties are spatially uniform across each galaxy. In the future, we will extend our lens models to overcome this limitation; here, we simply note that *abundance* variations – the chief uncertainty in measuring IMF variations spectroscopically – would contribute

negligibly to the gradient in $Y_\star(R)$ (for a 13.5 Gyr-old stellar population at solar metallicity, a change in abundance from $\left[\frac{\alpha/\text{Fe}}{\alpha/\text{Fe}_\odot}\right] = 0$ to $\left[\frac{\alpha/\text{Fe}}{\alpha/\text{Fe}_\odot}\right] = +0.4$ leads to changes in colour $\Delta(g-r) = 0.03$, $\Delta(r-i) = 0.01$, $\Delta(i-z) = 0.00$ according to the stellar population models of Conroy & van Dokkum, 2012a). This means that our lensing/dynamical inference is subject to a different degeneracy (IMF-age-metallicity) from stellar population studies (IMF-abundance). The fact that the only stellar population property that both these degeneracies have in common is the IMF adds strength to the argument that the variations being traced by both sets of studies are due to the IMF, but joint lensing/dynamics/stellar populations analyses are needed in order to confirm this hypothesis, and will be a subject of future work.

Finally, we note that a limitation of our inference on $Y_\star(R)$ is the power-law parameterisation that we impose. In contrast, previous constraints on the IMF (or Y_{tot}) have been able to make independent measurements at a number of different radii. It is necessary to assume parametric forms for the halo and stellar masses in order to disentangle them, but this may also be a source of bias; indeed, the reliability of our decomposition of the dark and luminous mass depends on the use of realistic density profiles for each. The identification of more motivated parametric forms for $Y_\star(R)$ using simulations will be an important next step.

8.4.2.2 A two-phase formation scenario

The Y_\star gradients we infer imply that the stellar population properties vary systematically across the galaxy. Though we cannot disentangle the contributions to these variations from stellar age, metallicity and the IMF, we note that our findings are qualitatively consistent with the suggestions of IMF gradients in ETGs that have been raised by recent stellar population modelling studies. Since the IMF is determined by the initial physical conditions in which a galaxy forms stars, variations in the IMF may indicate that different assembly processes dominate in the inner and outer regions of ETGs.

The two-phase scenario of ETG formation that has been developed to explain the rapid size evolution of these systems (e.g. Naab et al., 2009) seems to provide a natural interpretation for such gradients. In this picture, an ETG initially forms a compact core via rapid dissipational collapse, and then grows much more gradually by the accretion of lower-mass satellites at larger radii. If the first process gives rise to stellar populations with a bottom-heavy IMF – possibly due to the higher temperatures and densities of the star-forming regions, resulting in a smaller Jeans scale and so an excess of low-mass stars – and the objects accreted in the second phase themselves formed in more Milky-Way-like

conditions with Milky-Way-like IMFs, then this would lead to an IMF variation which is at least qualitatively consistent with the variations that have been inferred. We note that an age gradient might also be expected in this case, since the accreted stellar populations are likely to be younger than the ETG core. Together with further observational studies, more quantitative predictions from simulations are now needed to test this hypothesis more thoroughly.

8.4.3 The evolution of massive ETGs

We have found evidence that isolated, massive ETGs have cuspy haloes and heavy, radially-declining stellar-mass-to-light ratios. These results are consistent with a scenario in which the halo has been adiabatically contracted and the IMF becomes bottom-heavy at small galactocentric radii, as presented in the previous sections. We now attempt to draw these two results together to develop a coherent picture for the evolution of both the dark and the luminous mass.

Firstly, the steep central halo slopes suggest that these haloes have been contracted by the initial infall of gas. The high Υ_\star in the central regions then implies that this gas may have formed stellar populations in different physical conditions from the Milky Way, leading them to have a bottom-heavy IMF. If these conditions are determined (at least partly) by the depth of the central gravitational potential well in which the stellar populations are formed, then the relationship between the IMF mismatch and the inner halo slope may provide additional insights into the physics governing this stage of ETG assembly. In our small galaxy sample, we find tentative evidence that Υ_\star changes more steeply in more massive galaxies, but larger samples are needed before this relationship can be established conclusively.

Secondly, the large-scale environment of an ETG appears to be key to determining the halo structure in its inner regions. This may indicate that, though all haloes initially experience some contraction, those in denser environments are subsequently subject to stronger heating processes due to satellite accretion, which undoes some of the initial contraction. This raises the question of whether the strength of any $\Upsilon_\star(R)$ gradients should also exhibit an environmental dependence. That is, the higher rate of accretion events in BCGs relative to field galaxies – which leads to the expansion of their haloes – might also lead to steeper IMF gradients in these systems. On the other hand, the much greater spatial extent of the halo relative to the stellar mass may make it more sensitive to the large-scale environment. The role of AGN outflows in modifying the dark and luminous mass structure also remains unclear. Investigating the dependence of the IMF on galaxy

environment will be an important test of the evolutionary scenario put forward in this work.

8.5 Summary and conclusions

We have combined pixel-based strong lens modelling with Jeans dynamical modelling to construct models for the dark and luminous mass structures of 12 isolated ETGs at $z \sim 0.3$ and have reached the following conclusions.

1. Most of the sample have dark matter haloes which are centrally steeper than the NFW profile, such that the population is consistent with being drawn from a Gaussian distribution with mean $\gamma(R < R_e) = 1.66 \pm 0.11$. This is consistent with models of adiabatic contraction.
2. The cusped nature of these isolated ETGs is in contrast to results for more massive ETGs in denser environments, and presents evidence that the relative importance of different baryonic processes may depend on large-scale environment.
3. The majority of the sample have stellar-mass-to-light ratio gradients which decrease with increasing radius, such that the population is consistent with being drawn from a Gaussian distribution with mean $\alpha_{ML} = -0.22 \pm 0.08$. This is consistent with a scenario in which the IMF is bottom-heavy in the central regions and Milky-Way-like at large radii, and may reflect the two-phase assembly of massive ETGs.
4. The strength of the stellar-mass-to-light ratio gradients may increase with increasing projected central dark matter mass and halo slope, suggesting that the baryonic mass structure may depend on the depth of the central gravitational potential well.

Extending these methods to larger lens samples will allow a more thorough hierarchical analysis of the trends within the population of isolated ETGs and will be the subject of a future work.

CONCLUSIONS

This thesis has focussed on probing the dark and luminous mass structure of ETGs in order to improve our understanding of their evolution. In the Introduction, we presented three major questions for ETG growth, which we now reassess in the light of our results.

9.1 What is the structure of the dark halo in ETGs?

The small number of studies of ETG halo structure have found evidence that the dark halo in cluster-scale ETGs is expanded relative to the dark-matter-only prediction, whereas that in group-scale ETGs is mildly contracted (Newman et al., 2013, 2015); on the scale of isolated ETGs, no study so far has been able to break the degeneracy between the dark and luminous mass (Grillo, 2012; Sonnenfeld et al., 2012, 2015). It has therefore remained unclear whether the differences in the halo structure of cluster-scale and group-scale ETGs are tracing a real environmental trend, and further, independent studies on all mass scales have been needed.

In Chapters 2 and 3, we made initial steps to address this need by investigating the dark halo structure of the BCG M87, which we found to be centrally cored, consistent with the Newman et al. (2013) result; we also confirmed the robustness of this conclusion against the use of more flexible stellar mass models and improved stellar kinematics in Chapter 7. In Chapter 8, we then extended constraints on the halo structure to isolated strong lensing ETGs, which – unlike M87 – we found to be centrally contracted relative to dark-matter-only models. Taken together with the earlier results of Newman et al. (2013,

2015), the very different conclusions of our two studies suggest a picture in which the density of an ETG's environment plays a significant role in determining the inner halo structure. We have therefore suggested a simple scenario in which an initial, pristine ETG halo becomes contracted due to the infall of baryons as the luminous part of the galaxy forms (Blumenthal et al., 1986), and is subsequently expanded via dynamical heating as the ETG accretes lower-mass systems (El-Zant et al., 2004).

The scenario presented above would naturally reproduce the environmental dependence of the halo structure that we, and others, have inferred, and larger samples of ETGs in all environments must now be modelled in order to further test this hypothesis. It will also be important to investigate more realistic analytic profiles for both the dark halo and the stellar mass in order to assess the robustness of our disentanglements of these two components, and to make more direct, quantitative comparisons with simulations.

9.2 What is the nature of the IMF in ETGs?

Independent pieces of evidence from strong lensing, stellar dynamics and stellar population modelling agree that the IMF in ETGs appears to be bottom-heavy relative to that of the Milky Way (Auger et al., 2010b; van Dokkum & Conroy, 2010; Cappellari et al., 2012); however, more tentative evidence from stellar population modelling at optical wavelengths has recently been put forward to suggest that the IMF may vary radially within individual galaxies (Martín-Navarro et al., 2015a; van Dokkum et al., 2016), though the same conclusions have not been reached by studies of stellar populations at near-infrared wavelengths (Vaughan et al., 2016; Zieleniewski et al., 2017; Alton et al., 2017) or molecular gas dynamics (Davis & McDermid, 2017), and complementary constraints from strong lensing and larger-radius kinematics have so far been lacking. In Chapter 7, we provided the first constraints on stellar mass-to-light ratio variations using multiple dynamical tracers in M87. We found that its stellar mass-to-light ratio declines as a function of radius, with a gradient that cannot be explained by variations in the age, metallicity or dust extinction of its stellar populations with a spatially uniform IMF, but can be reproduced if the low-mass slope of the IMF is allowed to vary. Moreover, in Chapter 8, we found that the assumption that the stellar-mass-to-light ratios of our lensing ETGs are spatially constant leads to unrealistically cuspy dark haloes, but that a more flexible model in which the stellar-mass-to-light ratio follows a power law profile with radius alleviates this problem and that the majority of these ETGs also have radially declining stellar-mass-to-light ratios.

These findings seem to support previous results from stellar population studies at

optical wavelengths (Martín-Navarro et al., 2015a; van Dokkum et al., 2016), and are consistent with a scenario in which ETG formation happens in two distinct phases, with the rapid dissipational collapse of an initial compact core – with a bottom-heavy IMF – followed by the accretion of lower-mass systems – with Milky-Way-like IMFs – at larger radii. The next step will be to reconcile stellar population models in different wavelength regimes, and subsequently combine strong lensing and dynamics with stellar population models across long wavelength baselines in order to make firmer and more conclusive statements about the nature of the IMF in these systems. Again, closer comparison with simulations will also be important in order to investigate the astrophysical processes underlying these variations.

9.3 What are the physical mechanisms driving ETG growth?

The strong evolution in ETG compactness since $z \sim 2$ has led to a picture in which the growth of these systems is driven by successive minor mergers and accretion events. However, observational limitations have made it difficult to test the structures of individual systems against predictions of minor merger models. In Chapter 5, we used strong gravitational lensing to super-resolve a sample of 13 massive, compact ETGs at $z \sim 0.6$, which we found to have core+envelope structures consistent with the predictions of minor merger models. We then constructed the fundamental plane for these systems in Chapter 6 to show that, despite their low central dark matter fractions, their fundamental plane is tilted relative to both the virial plane and the nominal fundamental plane, indicating that ETGs are non-homologous in both their dark and luminous mass structures – thus indicating some scatter in the relative importance of different growth mechanisms. In the future, it may be possible to connect this non-homology to the underlying evolutionary processes at work in these systems.

Finally, in Chapter 4 we constructed scaling relations for a sample of cluster ETGs at similar redshifts to our compact ETG sample, and used these to show that cluster ETGs appear to have undergone extremely little structural evolution since $z \sim 0.6$, especially compared with their isolated counterparts. We suggested that the implication of this result – that galaxies in dense environments experience accelerated growth at earlier times – could be a result of the higher merger rates that are achievable for such systems, and so interpret this result as further evidence that hierarchical processes dominate the growth of ETGs.

9.4 A coherent view of ETG evolution

We are now in a position to combine these different results and so formulate a coherent picture of ETG evolution. Firstly, we have found that the inner slope of the dark matter halo depends on the large-scale environment, such that ETGs at the centres of clusters have expanded haloes whereas the haloes of isolated ETGs are strongly contracted. Second, we have found possible signatures of minor mergers in the luminous structures of anomalously compact ETGs. Thirdly, we have presented evidence that the growth rate of ETGs at early times is accelerated in dense environments. Finally, we have shown that the stellar-mass-to-light ratios of ETGs vary radially, consistent with a picture in which these systems have bottom-heavy IMFs in their central regions but Milky-Way-like IMFs at larger radii.

These key results can be explained coherently – at least on a qualitative level – if ETG assembly and evolution is dominated by hierarchical processes. In this paradigm, an ETG forms a compact core at high redshifts – leading to the compact ‘red nuggets’ that are observed at $z \sim 2$, and in physical conditions which lead the stellar populations to have bottom-heavy IMFs. Moreover, the infall of this baryonic material causes the dark halo to contract. Subsequently, the ETG grows by the accretion of many lower-mass systems. Since these accreted satellites have lower stellar masses and so did not form in the same – possibly extreme – conditions as the ETG core, they have Milky-Way-like IMFs, such that their accretion onto the core creates the stellar-mass-to-light ratio gradients that we observe. Furthermore, the fact that this growth is dominated by minor mergers and accretion means that (a) ETGs in clusters grow more rapidly than those in the field, and (b) ETGs in cluster centres experience more dynamical heating due to the frictional effects of infalling systems, such that their dark haloes become more expanded. This, then, also gives rise to the environmental dependence of the inner halo structure that we infer.

9.5 Future directions

Given the simple picture presented in the previous section, the task is now to test and develop it more rigorously. The first important step here will be to extend the methods of Chapter 8 to significantly larger galaxy samples. There exist ~ 150 systems in the Sloan Lens ACS Survey (SLACS; Bolton et al., 2008), BOSS Emission Line Lens Survey (BELLS; Brownstein et al., 2012) and Strong Lensing Legacy Survey (SL2S; Gavazzi et al., 2012) with sufficiently high-quality imaging and sufficiently simple lensing configurations for our techniques; moreover, these systems span a ~ 5 times wider range of stellar masses than the EELs, and so will allow a much more conclusive investigation of trends within

the ETG population. Extending our methods to group- and cluster-scale lenses, with more sophisticated dynamical modelling out to large radii, will also tighten constraints on these scales and so allow firmer conclusions on the possibility that ETG halo structure varies with environment. In tandem with this, it will be important to work closely with simulations to develop more realistic analytic profiles with which to describe the dark and luminous mass distributions, so as to reduce systematic uncertainties in our inference on both.

On the other hand, the work in this thesis represents the latest step in the development of methods, initiated more than a decade ago, to infer ETG mass structure through the combination of strong lensing and stellar dynamics. These probes can constrain the total stellar mass and therefore the IMF *normalisation*; more recently, however, it has become clear that careful stellar population modelling of ETG spectra can provide complementary information on the IMF *slope*. The next key step in this generation of lensing+dynamical studies, then, will be the incorporation of stellar population modelling methods in a self-consistent way, which will allow direct inference on both IMF properties simultaneously. Moreover, the addition of stellar population modelling constraints on the stellar mass in these systems will allow a more robust inference on the halo structure and the use of more flexible halo models. We are currently embarking on a pilot study of this nature by combining MUSE spectra and HST imaging of the lowest-redshift lens, ESOJ1343-3810, and plan to extend our investigations to higher redshifts in the future.

Finally, we have demonstrated the utility of strong lensing magnification as a probe of the *luminous* structure of compact ETGs and therefore of their channels of growth. Identifying higher-redshift EELs spectroscopically is difficult, as the wavelengths of interest become shifted into sky-dominated regions of the spectrum; however, recent deep imaging surveys such as the Dark Energy Survey (DES; Dark Energy Survey Collaboration et al., 2016) have motivated the development of sophisticated photometric lens searches, which may represent a way forward here. Identifying and modelling lensed compact ETGs in higher-redshift regimes will allow us to extend our understanding of the luminous structure and size growth of compact ETGs to earlier stages in their evolution.

The coming years will see the launch of a number of next-generation surveys and instruments such as LSST, eBOSS and Euclid, which are forecast to discover galaxy-scale strong lenses in their thousands (Collett, 2015). The methods developed in this thesis – in terms of understanding both the mass and the luminous structure of ETGs – will therefore gain significant statistical power when these large samples become available.



APPENDIX A: ANISOTROPY KERNEL FOR THE SCALED OSIPKOV-MERRITT PROFILE

For the anisotropy profile of Equation 3.21, the integrating factor defined by Equation 3.18 is

$$(A.1) \quad f(r) = f(r_i) \left(\frac{r^2 + r_a^2}{r_i^2 + r_a^2} \right)^{\beta_\infty}.$$

The kernel of Equation 3.20 is then

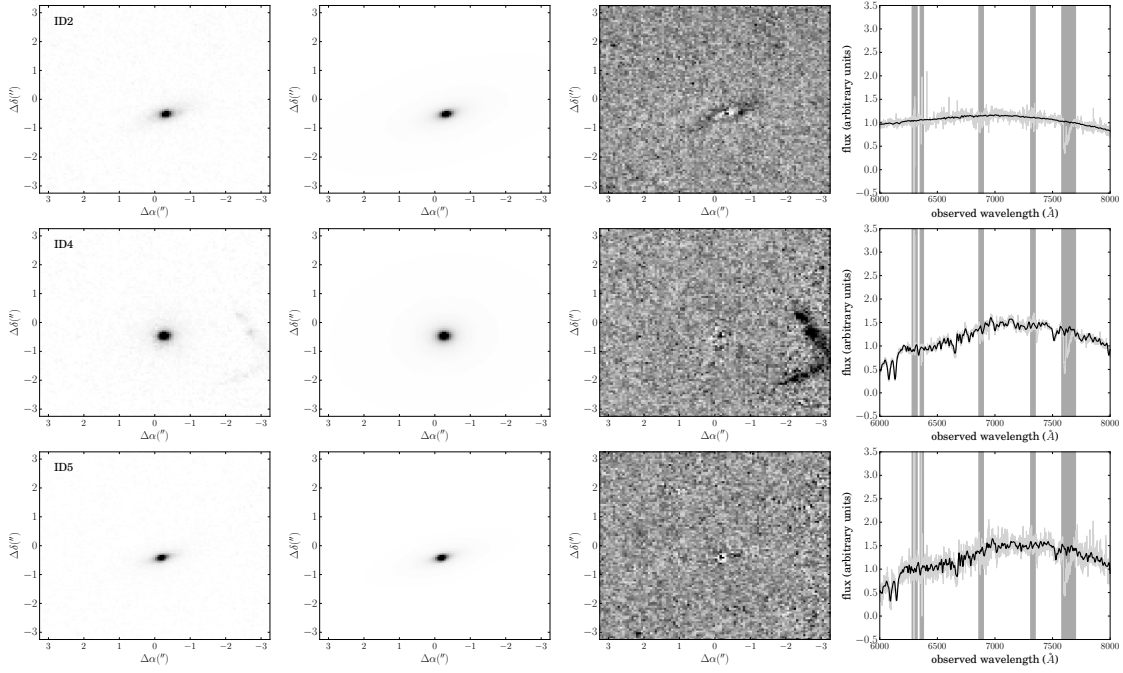
$$(A.2) \quad K_\beta(R, r) = \frac{\sqrt{r^2 - R^2}}{r} \left(1 + \beta_\infty \left(\frac{r^2 + r_a^2}{R^2 + r_a^2} \right)^{\beta_\infty} \left[{}_2F_1\left(\frac{1}{2}, \beta_\infty; \frac{3}{2}; z\right) - \frac{3R^2 + 2r_a^2}{R^2 + r_a^2} {}_2F_1\left(\frac{1}{2}, \beta_\infty + 1; \frac{3}{2}; z\right) \right] \right)$$

where

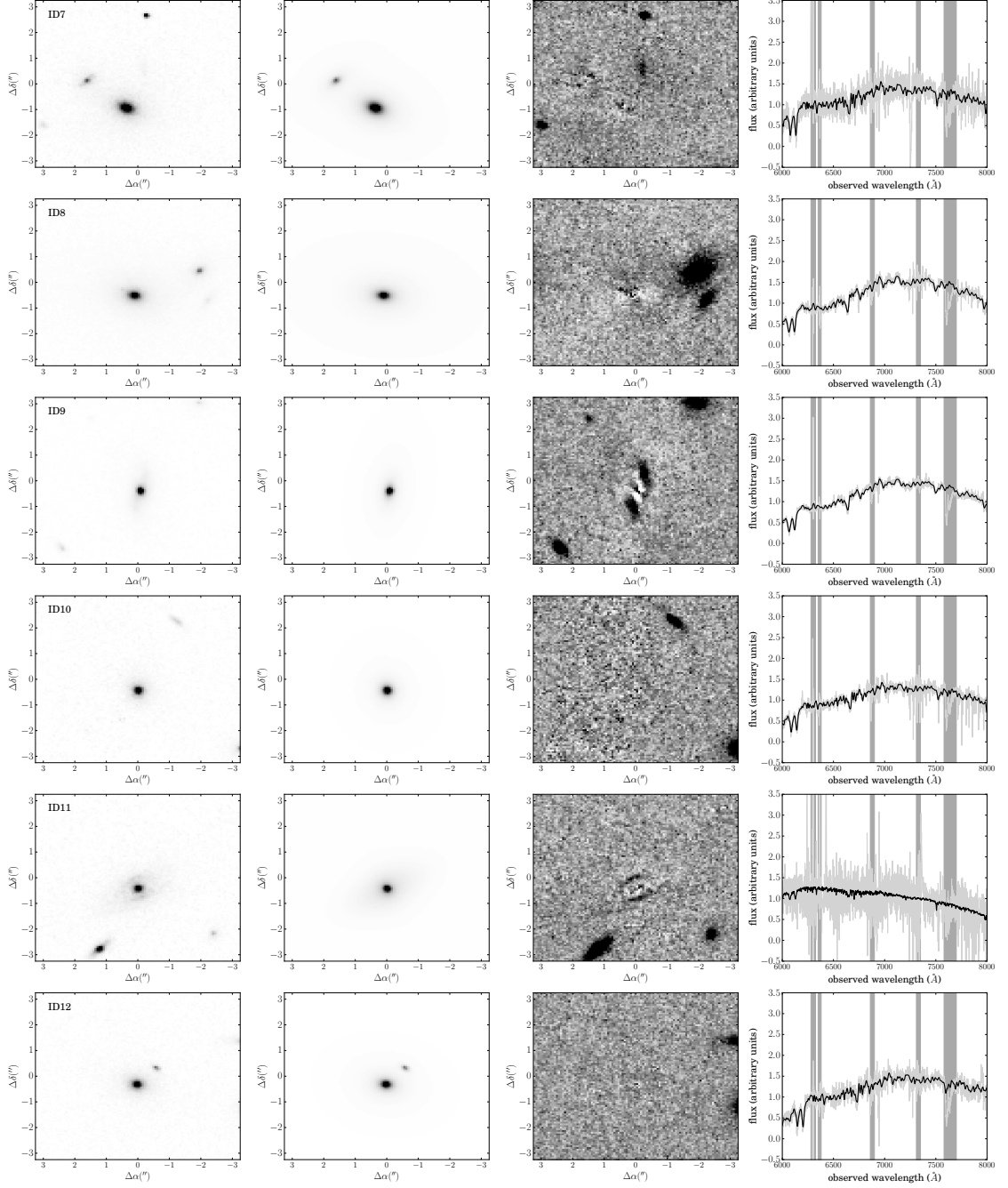
$$(A.3) \quad z = \frac{R^2 - r^2}{R^2 + r_a^2}.$$

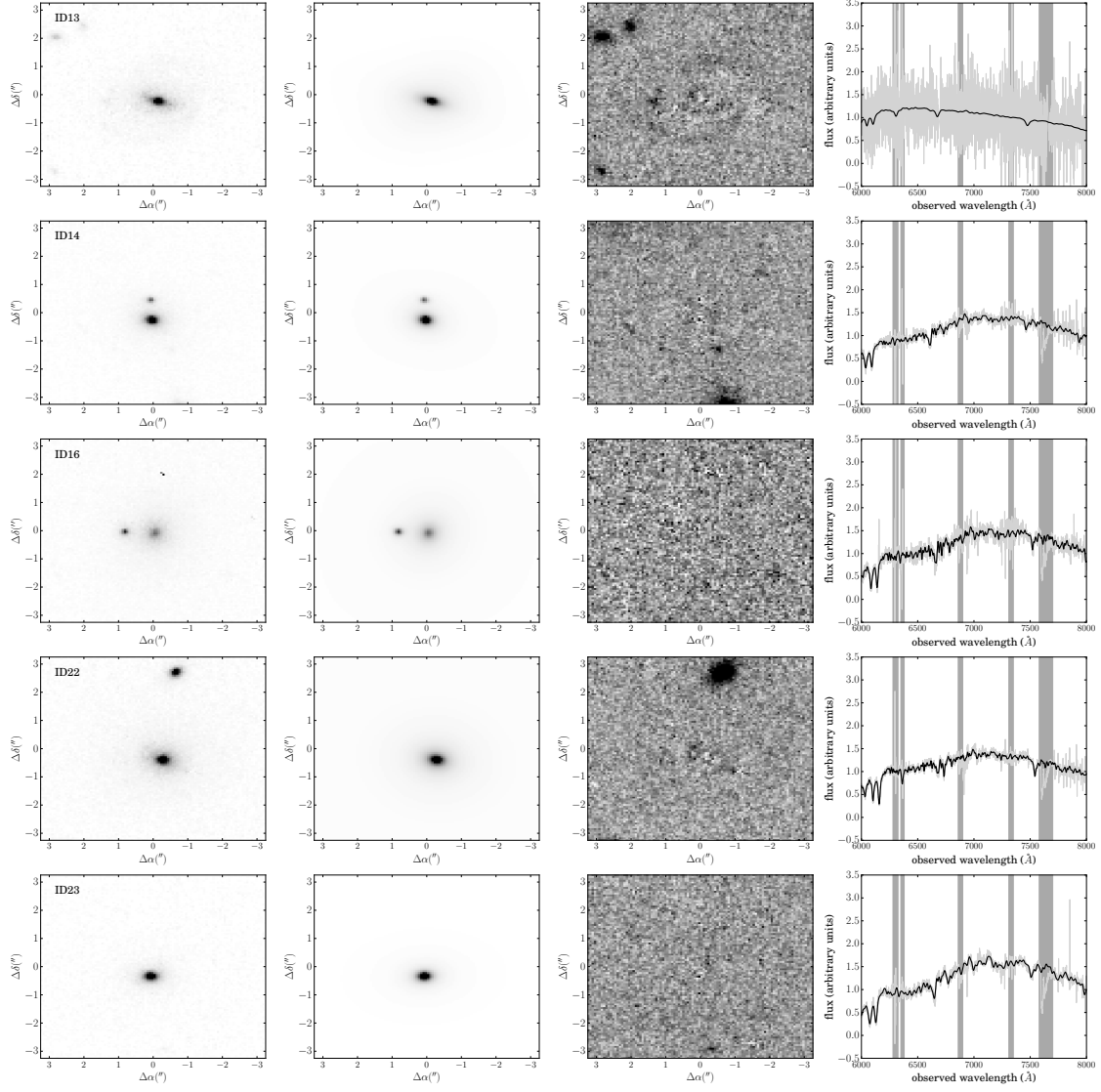
APPENDIX B

APPENDIX B: PHOTOMETRY AND SPECTRA OF THE FINAL MACSJ0717 GALAXY SAMPLE



APPENDIX B. INDIVIDUAL MACSJ0717 GALAXIES





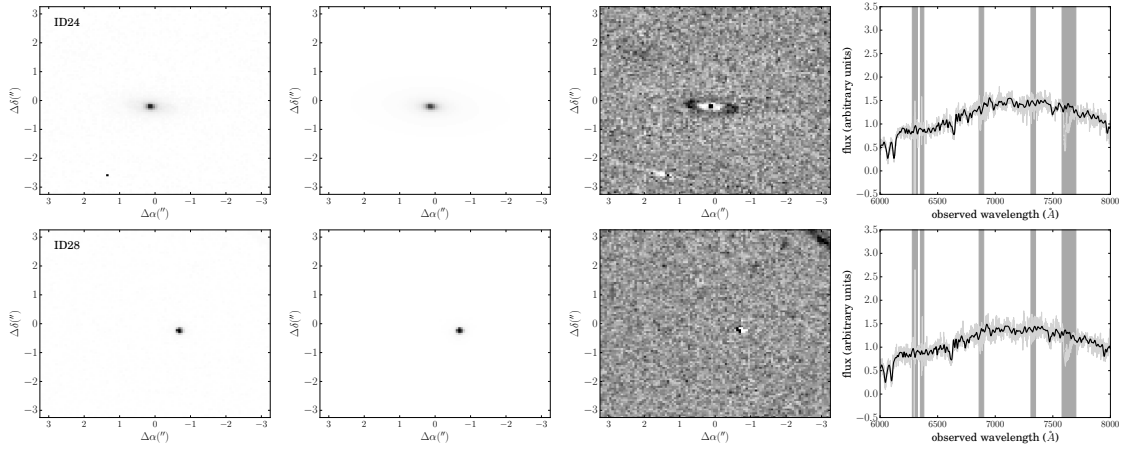


Figure B.1: For each object in our final MACSJ0717 sample, from left to right: HST/ACS F625W data; model; signal-to-noise residuals; fitted spectrum. Image cutouts are $6.5''$ on a side; spectra show the extracted spectrum in grey and the model in black, with masked regions in dark grey.

APPENDIX C

APPENDIX C: NOTES ON INDIVIDUAL EELS

As explained in Section 4, some EELs presented unusual or interesting features or were not well described by 2C models. We summarise these systems here, and present pixellated source reconstructions for all thirteen systems in Figure A1.

1. While the source in **J0837** appears fairly simple in the K' band, the HST data reveal a clear dip across the middle of both arcs. Since this appears in both images, it is much more likely to be related to the source as opposed to any perturbations in the lensing mass (e.g. Koopmans, 2005). We therefore assume this dip in the surface brightness to be due to a dust lane in the source, and model it using a second Sérsic component which we require to have a negative amplitude. This significantly improves the source model, and suggests that this galaxy may have undergone a recent merger. Our pixellated reconstruction – shown in Figure D.1 – also recovers this dust lane.
2. Neither the 1C nor the 2C model for **J1125** was able to fully account for the brightness of the lower arc of the source. This is especially apparent in the I band residual image, and indicates that even a double-Sérsic profile model may not be a good description of the source in this case. Moreover, the bulge component of the 2C model has an extremely small effective radius $R_e = 0.24$ kpc and a high surface brightness (despite its small size, the bulge-to-total ratio in the I band is still $B/T_I = 0.71$); the more ‘extended’ component is also quite compact at $R_e = 1.49$ kpc. This suggests a bright compact source such as an AGN. Our pixellated models similarly fail to

fully describe the brightest pixels in the arc; since we optimise these models for a regularisation which is constant across the image, this also seems to suggest the presence of an extremely compact central component which our regularisation may be smoothing away. It is also possible that the central component may be offset from the more extended one, either physically or due to dust obscuration. This is apparent in the slight asymmetry of the pixellated source, and may be an additional reason why our concentric parametric models cannot fully describe the data here. Indeed, when we relax this condition in our parametric model, the two source components do become offset by $\sim 1.3\text{kpc}$, though the remaining properties of both source and lens light profiles and the lensing mass profile remain consistent with those of the concentric model.

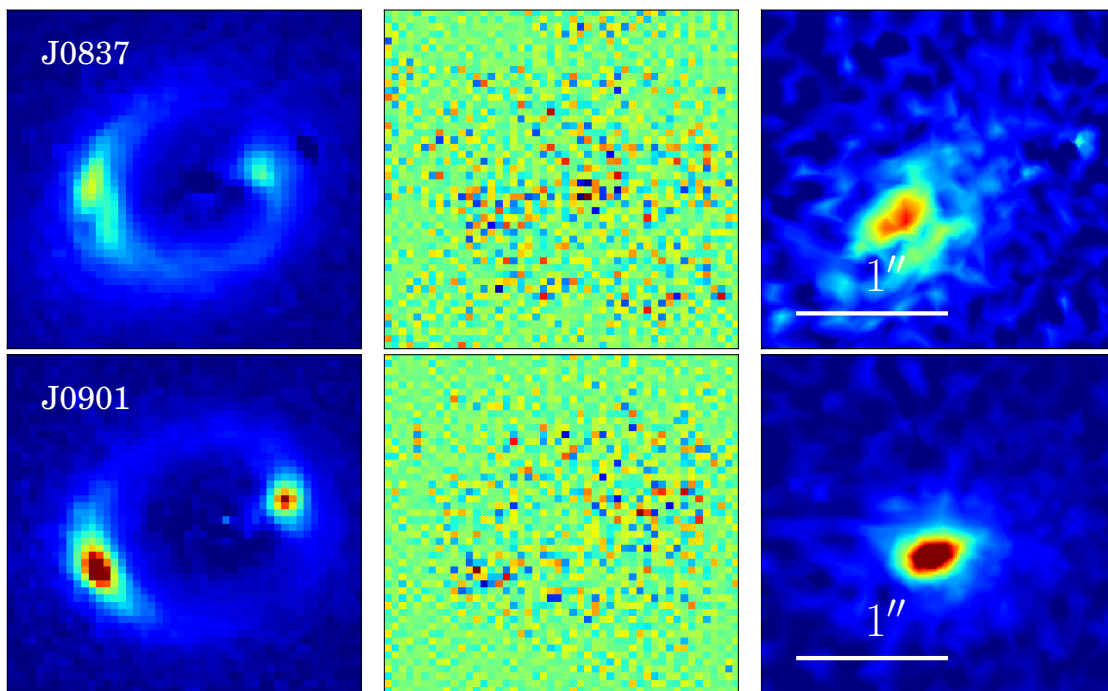
3. As a check on our inference on the source structure, we note that the K' band image of **J1347** has been modelled previously by Auger et al. (2011), and we compare our results for this object with the model reported in that study. As here, Auger et al. (2011) also find that a two-component fit is necessary to accurately model the surface brightness distribution, and that the inferred size of the source significantly increases when the second component is included. On the other hand, the total radius of our 2C model is $R_e = 3.96 \pm 0.33\text{ kpc}$, which is significantly larger than their 1.1 kpc , and this difference is also seen in the inferred magnification (compare our $\mu = 5.09$ with their $\mu = 12$). This difference may be driven by differences imposed by the models or by the data, as the current analysis also includes the ACS optical data. Also, Auger et al. (2011) required the bulge component to follow a de Vaucouleurs profile with $n = 4$, whereas we left this as a free parameter and found $n = 7.86$, and this then has repercussions for the structure of the envelope component: indeed, Auger et al. (2011) finds a Sérsic index of $n_{env} = 0.6$ which is substantially smaller than our $n_{env} = 1.44$. We also infer a power-law mass profile for the lensing galaxy with $\eta = 1.23 \pm 0.01$, which is significantly steeper than the SIE that was assumed in the earlier work.
4. While the prominent disk in **J1446** does not appear to be lensed and therefore seems at first glance to be associated with the lens galaxy, we find that 1C models with a single source component and two lens galaxy components (in which the second is highly flattened) are unable to provide a good description of the data. Further, close examination of the disk and the lens galaxy bulge reveals that the bulge is in fact offset from the centre of the disk by $\sim 0.1\text{ arcsec}$. When we then create 2C models for this system, we find that the second source component becomes highly flattened and the model provides a very good description of the data. We are therefore led to the

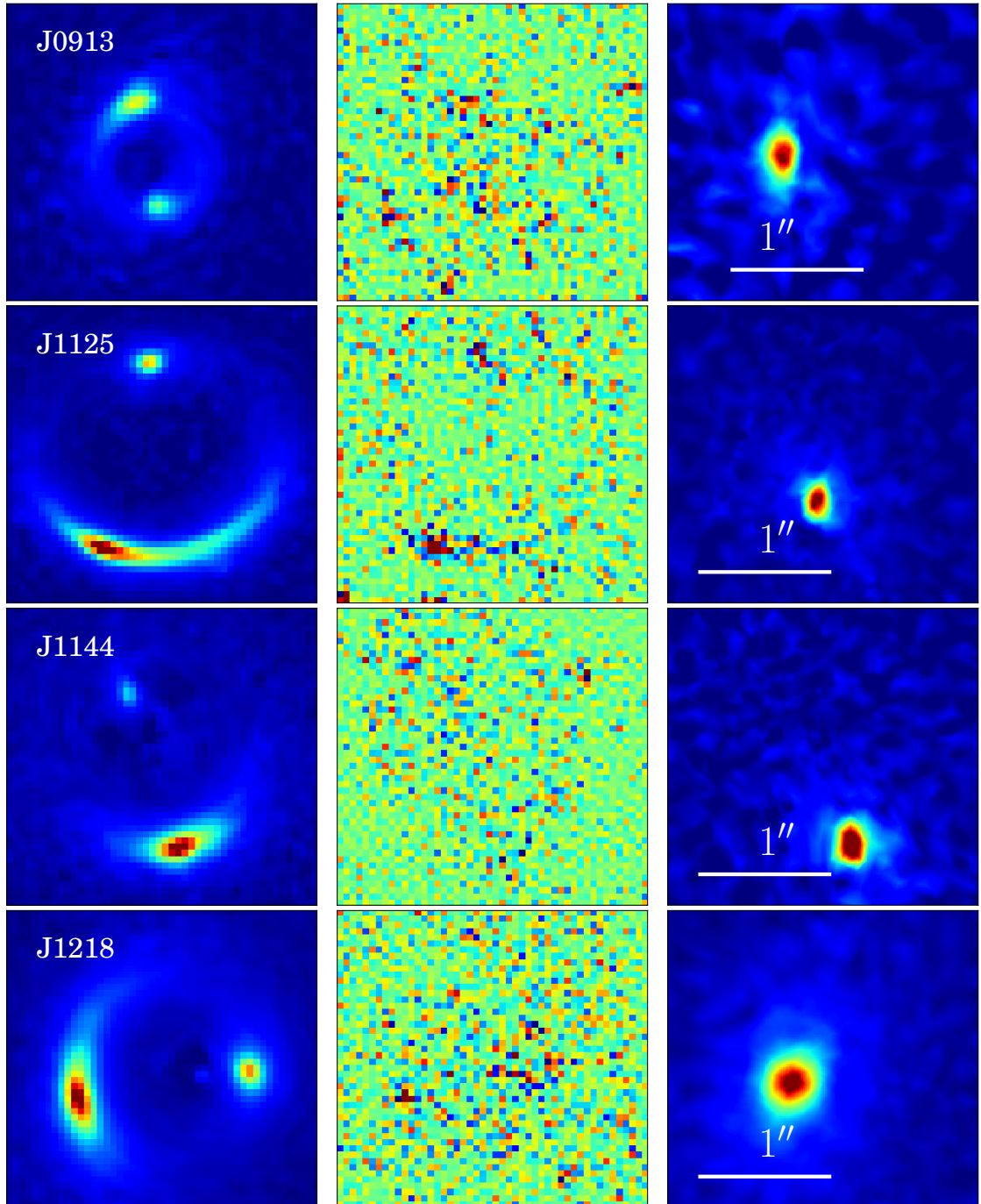
somewhat surprising conclusion that the disk is in fact associated with the source galaxy. At source redshift $z_s = 0.58$, the physical size of the disk is actually rather small at $R_e = 1.69 \pm 0.02$ kpc, but because it extends beyond the Einstein radius of the lens, the tips of the disk are not lensed and retain their distinct disk-like structure. The fact that this galaxy is clearly disk-like is interesting in light of the various scenarios put forward for red nugget growth and the finding by e.g. Stockton et al. (2014) and Hsu et al. (2014) of a high fraction of flattened galaxies in their moderate-redshift red nugget samples (as discussed in Section 6).

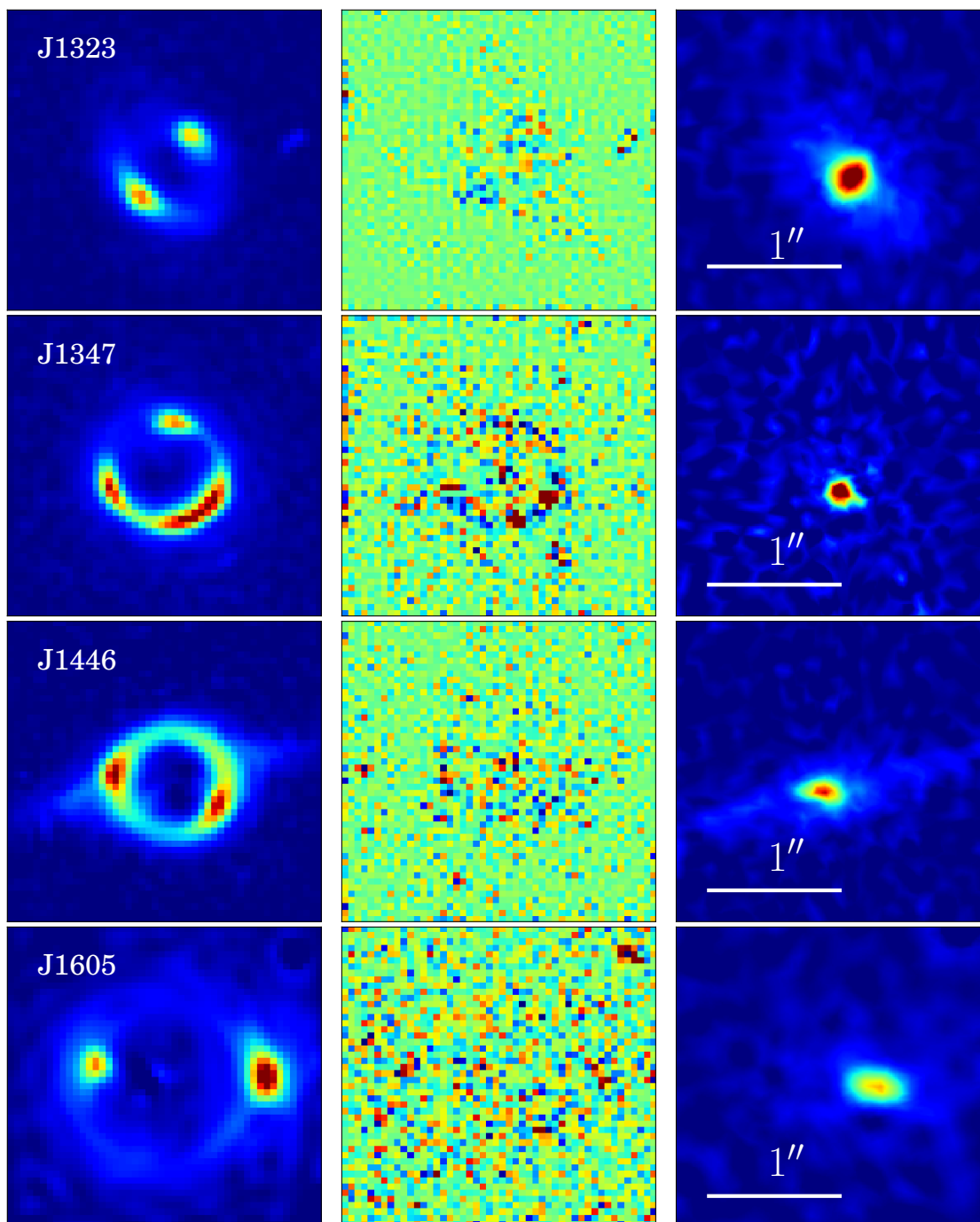
5. The source in **J1606** also exhibits a clear disk, although in this case it is almost totally lensed. Our 1C model for this system is really just a model for the bulge component and therefore provides a poor overall fit to the data; for our 2C model, we find that neither a highly flattened Sérsic nor an exponential disk profile can provide entirely satisfactory fits to the disk component, and we therefore implement the second source component as a boxy bulge, with a highly flattened Sérsic profile and circularised radial coordinate given by $r^c = (qx)^c + (y/q)^c$ where c is a free parameter in the model, with $c < 2$ indicating a diskiness and $c > 2$ indicating boxiness. We find $c = 3.44 \pm 0.20$, implying that the source in this system has a strong bar-like central surface brightness distribution.
6. While it is straightforward to find a good model for the V band image of **J1619** – where the signal-to-noise ratio is lowest – models which describe both the V and the I bands tend to leave unsatisfactory residuals in both filters, with an undersubtracted ring of flux at the Einstein radius and a slightly oversubtracted bulge component. Our pixellated source reconstruction indicates a significant asymmetry in the source which may explain this as a limitation of our Sérsic models. On the other hand, the pixellated model also has poor residuals, which suggests that the mass model may be at fault. For instance, there may also be a faint or dark perturber along the line of sight which our model does not include.

APPENDIX D

APPENDIX D: PIXELLATED RECONSTRUCTIONS OF THE EELS







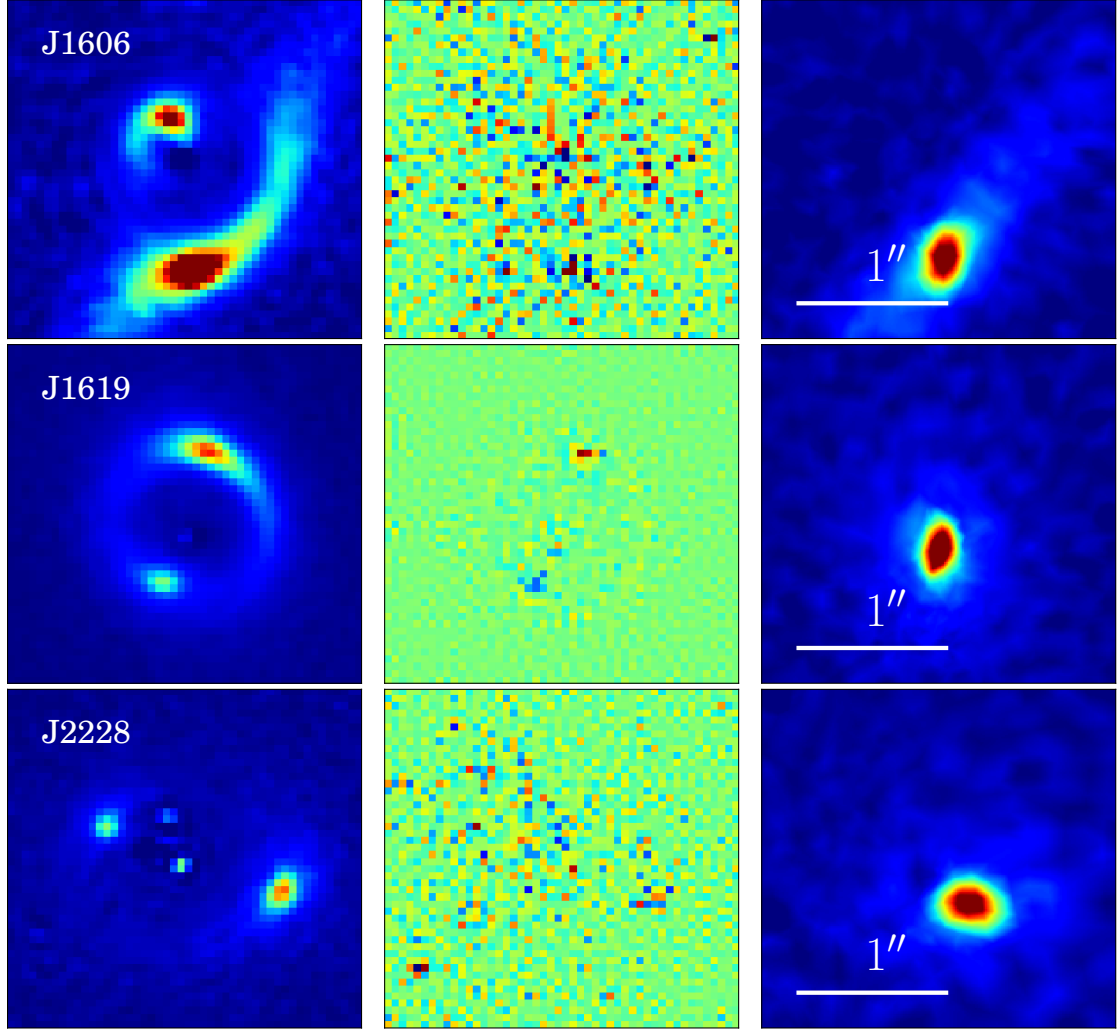


Figure D.1: Pixellated reconstructions for the thirteen EELs analysed here. From left to right, we show the V-band image, the signal-to-noise residuals and the reconstructed source. Note that these are not fitted models, but reconstructions of the source based on the lens models inferred using parametric source models. These reconstructions generally confirm that the sources are smooth, though they also recover the dust lane feature in J0837 and the disk features in J1446 and J1606.

APPENDIX E: INFERENCE ON M87'S MASS STRUCTURE FROM DYNAMICS AND PHOTOMETRY

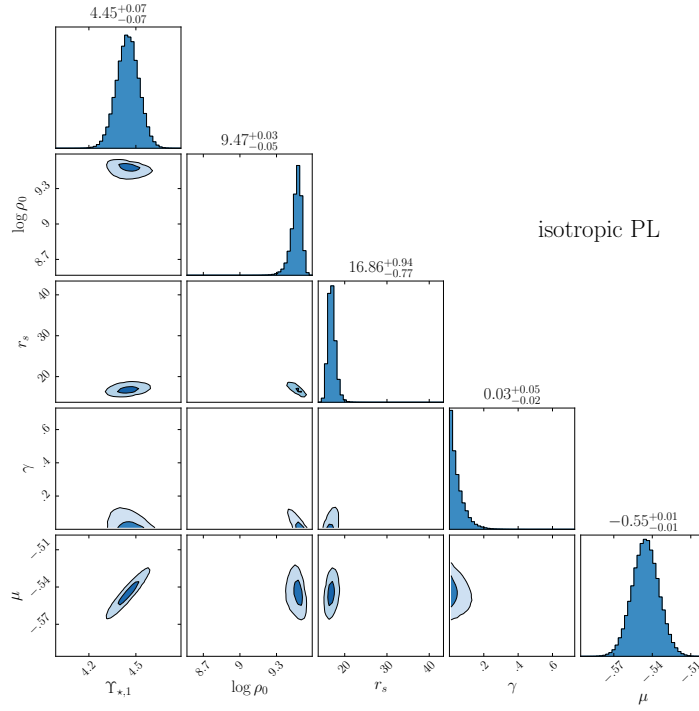


Figure E.1: Inference on the dynamical model under assumptions of isotropy for all tracer populations and the PL Y_\star model. Units are the same as in Table 1.

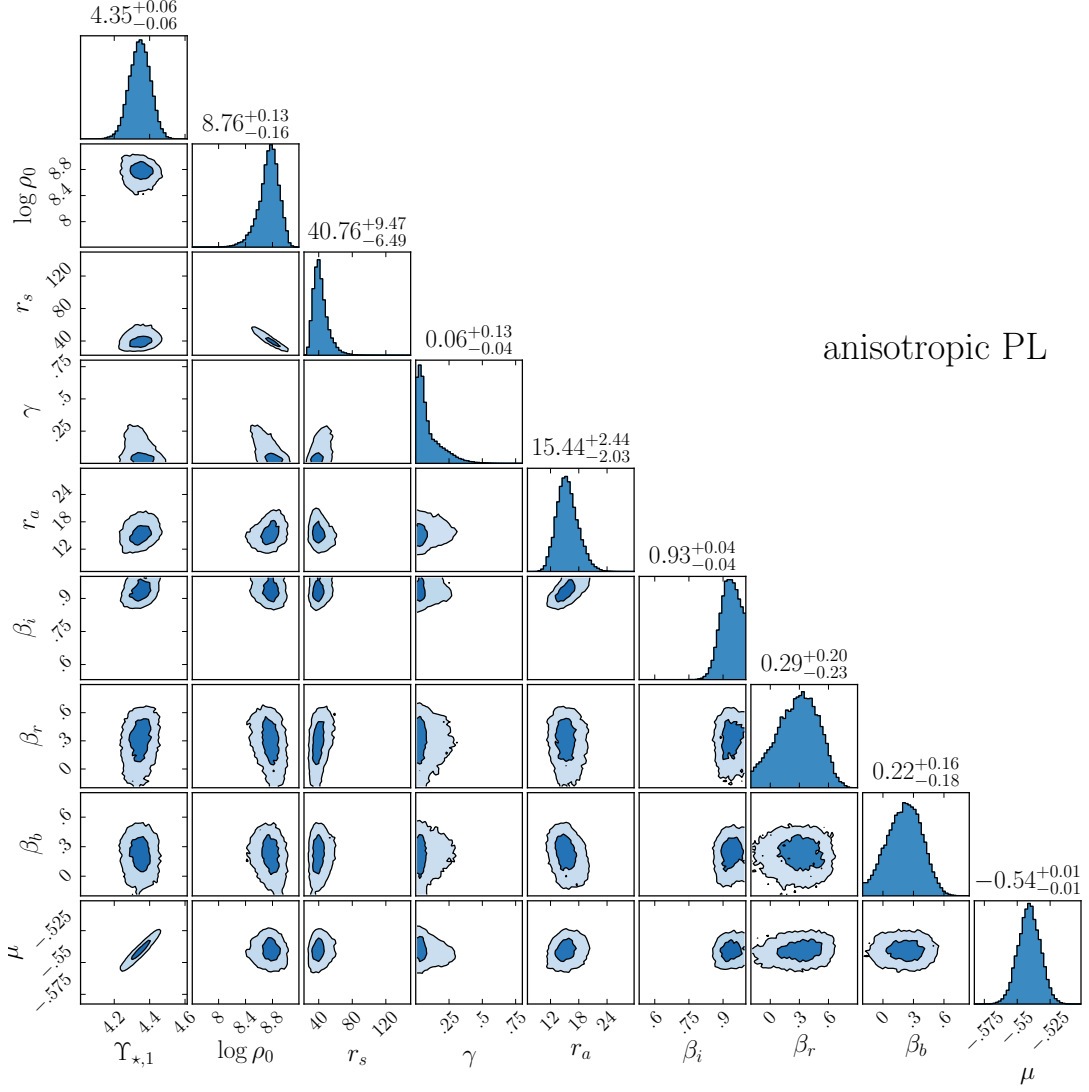


Figure E.2: Inference on the dynamical model under assumptions of anisotropy for all tracer populations and the PL Υ_* model, as detailed in Section 3. Units are the same as in Table 1.

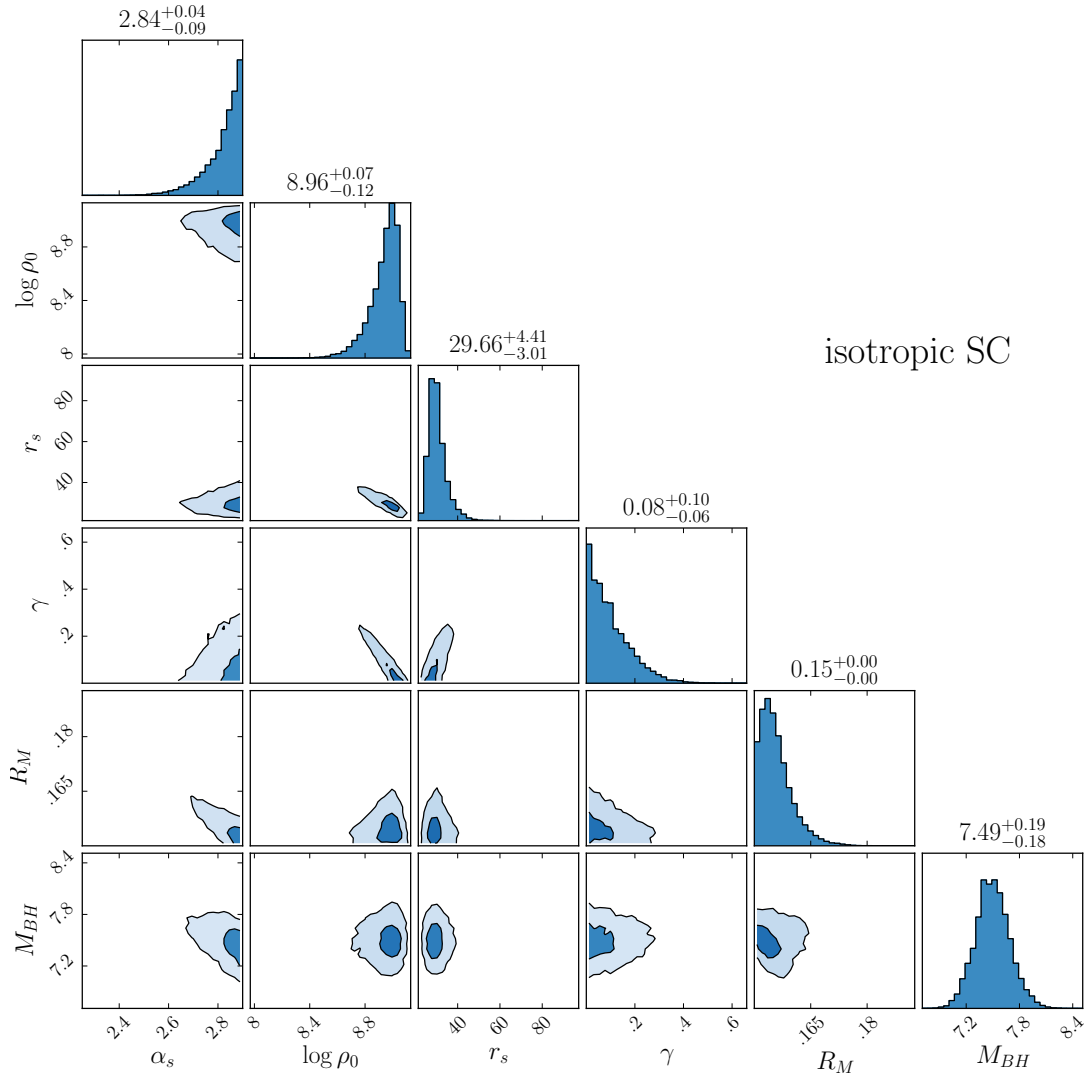


Figure E.3: Inference on the dynamical model under assumptions of isotropy for all tracer populations and the SC Y_\star model. Units are the same as in Table 1.

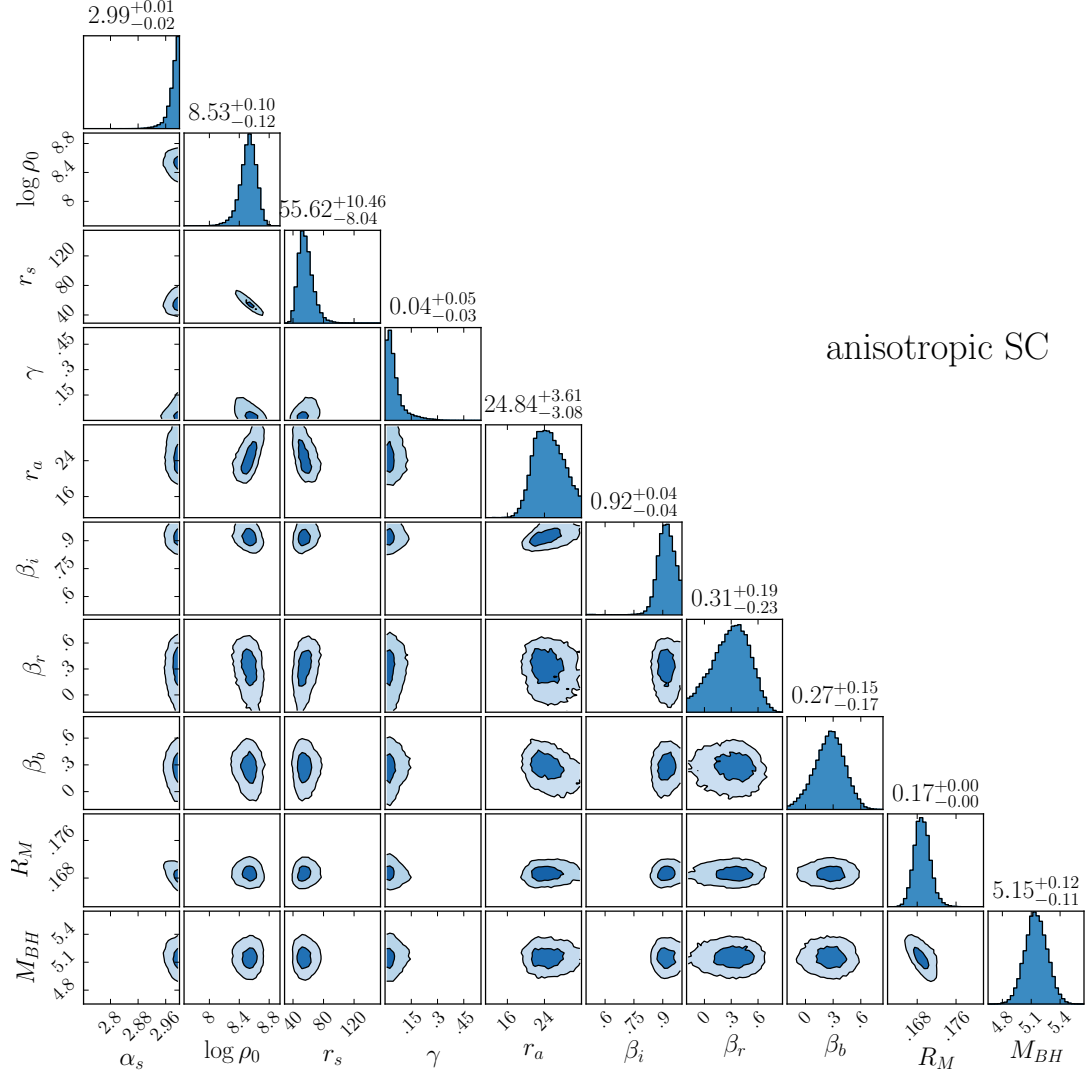


Figure E.4: Inference on the dynamical model under assumptions of anisotropy for all tracer populations and the SC Y_* model, as detailed in Section 3. Units are the same as in Table 1.

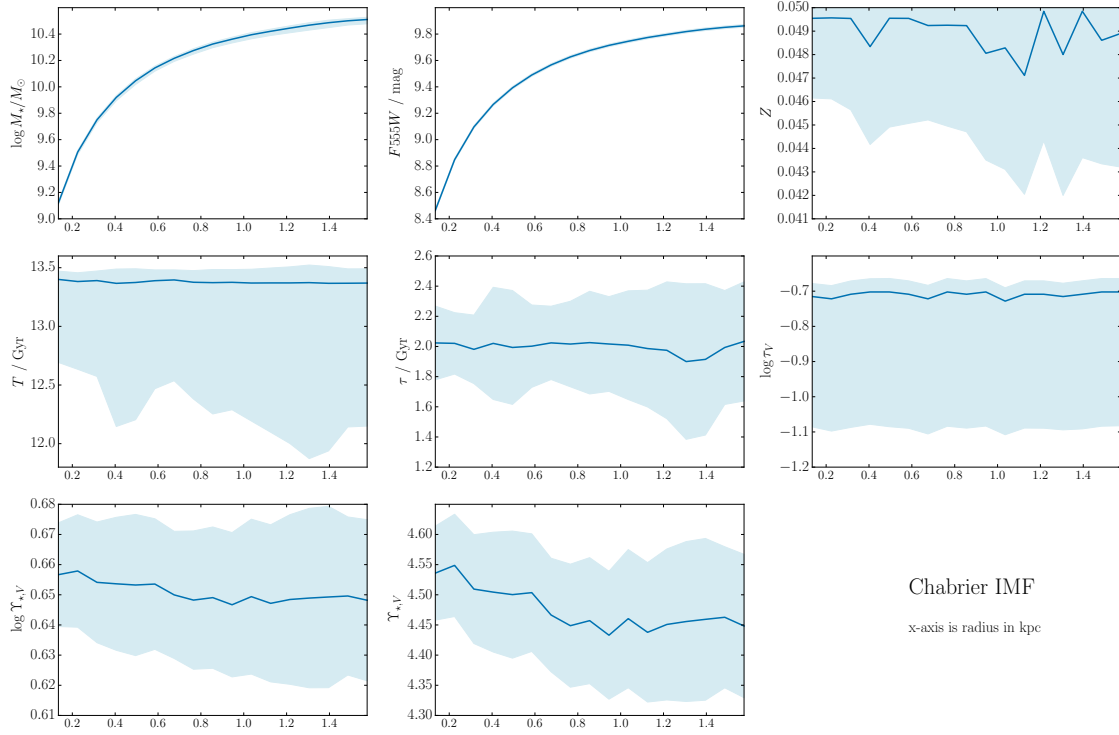


Figure E.5: Inference on stellar population properties, modelling high-resolution 11-band HST photometry using the models of BC03. Allowing for gradients in all parameters except the IMF allows only very weak gradients in the stellar mass-to-light ratio, suggesting that the main cause of the stellar mass-to-light ratio gradient that we infer dynamically may be IMF variations. This Figure shows our inference assuming a Chabrier IMF; our conclusions are qualitatively the same when a Salpeter IMF is assumed instead.

BIBLIOGRAPHY

- Abadi, M. G., Navarro, J. F., Fardal, M., Babul, A., & Steinmetz, M. 2010, MNRAS, 407, 435
- Dark Energy Survey Collaboration, Abbott, T., Abdalla, F. B., et al. 2016, MNRAS, 460, 1270
- Agnello A., Evans N. W., Romanowsky A. J., Brodie J. P., 2014, MNRAS, 442, 3299
- Alton, P. D., Smith, R. J., & Lucey, J. R. 2017, MNRAS, 468, 1594
- Ashman, K. M., & Zepf, S. E. 1992, ApJ, 384, 50
- Atkinson, A. M., Abraham, R. G., & Ferguson, A. M. N. 2013, ApJ, 765, 28
- Auger, M. W., Treu, T., Bolton, A. S., et al. 2009, ApJ, 705, 1099
- Auger, M. W., Treu, T., Bolton, A. S., et al. 2010, ApJ, 724, 511
- Auger, M. W., Treu, T., Gavazzi, R., et al. 2010, ApJL, 721, L163
- Auger, M. W., Treu, T., Brewer, B. J., & Marshall, P. J. 2011, MNRAS, 411, L6
- Bacon, R., Copin, Y., Monnet, G. et al. 2001, MNRAS, 326, 23
- Barkana, R. 1998, ApJ, 502, 531
- Barnabè, M., Czoske, O., Koopmans, L. V. E., et al. 2009, MNRAS, 399, 21
- Barnabè, M., Czoske, O., Koopmans, L. V. E., Treu, T., & Bolton, A. S. 2011, MNRAS, 415, 2215
- Barnabè, M., Spiniello, C., Koopmans, L. V. E., et al. 2013, MNRAS, 436, 253
- Barro, G., Faber, S. M., Pérez-González, P. G., et al. 2013, ApJ, 765, 104
- Bastian, N., Covey, K. R., & Meyer, M. R. 2010, ARA&A, 48, 339

- Baum W. A., 1955, *Publ. Astron. Soc. Pacific*, 67, 328
- Baum W. A., 1959, *PASP*, 71, 106
- Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, *ApJ*, 717, 379
- Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, *ApJ*, 770, 57
- Belli, S., Newman, A. B., & Ellis, R. S. 2014, *ApJ*, 783, 117
- Bender R., Ziegler B., Bruzual G., 1996, *ApJ*, 463
- Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393
- Bezanson, R., van Dokkum, P. G., Tal, T., et al. 2009, *ApJ*, 697, 1290
- Binney, J., & Tremaine, S. 2008, *Galactic Dynamics: Second Edition*, by James Binney and Scott Tremaine. ISBN 978-0-691-13026-2 (HB). Published by Princeton University Press, Princeton, NJ USA, 2008.,
- Blakeslee J. P., Jordán A., Mei S., Côté P., Ferrarese L., Infante L., Peng E. W., Tonry J. L., West M. J., 2009, *ApJ*, 694, 556
- Blumenthal, G. R., Faber, S. M., Flores, R., & Primack, J. R. 1986, *ApJ*, 301, 27
- Bolton, A. S., Burles, S., Koopmans, L. V. E., Treu, T., & Moustakas, L. A. 2006, *ApJ*, 638, 703
- Bolton, A. S., Burles, S., Treu, T., Koopmans, L. V. E., & Moustakas, L. A. 2007, *ApJL*, 665, L105
- Bolton, A. S., Treu, T., Koopmans, L. V. E., et al. 2008, *ApJ*, 684, 248-259
- Bower R. G., Kodama T., Terlevich A., 1998, *MNRAS*, 299, 1193
- Bower R. G., Lucey J. R., Ellis R. S., 1992, *MNRAS*, 254, 601
- Brown, M. J. I., Moustakas, J., Smith, J.-D. T., et al. 2014, *ApJS*, 212, 18
- Brownstein, J. R., Bolton, A. S., Schlegel, D. J., et al. 2012, *ApJ*, 744, 41
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Caminha, G. B., Grillo, C., Rosati, P., et al. 2015, *arXiv:1512.04555*

- Cappellari, M., & Emsellem, E. 2004, *PASP*, 116, 138
- Cappellari, M., Bacon, R., Bureau, M., et al. 2006, *MNRAS*, 366, 1126
- Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011, *MNRAS*, 413, 813
- Cappellari, M., McDermid, R. M., Alatalo, K., et al. 2012, *Nature*, 484, 485
- Cappellari, M., Scott, N., Alatalo, K., et al. 2013, *MNRAS*, 432, 1709
- Cappellari, M., McDermid, R. M., Alatalo, K., et al. 2013, *MNRAS*, 432, 1862
- Cappellari, M. 2016, *ARA&A*, 54, 597C
- Carollo, C. M., Bschorr, T. J., Renzini, A., et al. 2013, *ApJ*, 773, 112
- Chabrier, G. 2003, *PASP*, 115, 763
- Chakrabarty D., 2007, *MNRAS*, 377, 30
- Charlot, S., & Fall, S. M. 2000, *ApJ*, 539, 718
- Ciotti, L., Lanzoni, B., & Renzini, A. 1996, *MNRAS*, 282, 1
- Collett, T. E. 2015, *ApJ*, 811, 20
- Colín, P., Klypin, A. A., & Kravtsov, A. V. 2000, *ApJ*, 539, 561
- Conroy, C., Gunn, J. E., & White, M. 2009, *ApJ*, 699, 486
- Conroy, C., & Gunn, J. E. 2010, *ApJ*, 712, 833
- Conroy, C., & van Dokkum, P. 2012, *ApJ*, 747, 69
- Conroy, C., & van Dokkum, P. G. 2012, *ApJ*, 760, 71
- Côté P., McLaughlin D. E., Hanes D. A., Bridges T. J., Geisler D., Merritt D., Hesser J. E., Harris G. L. H., Lee M. G., 2001, *ApJ*, 559, 828
- Côté P., Marzke R. O., West M. J., 1998, *ApJ*, 501, 554
- Daddi, E., Renzini, A., Pirzkal, N., et al. 2005, *ApJ*, 626, 680
- Damjanov, I., McCarthy, P. J., Abraham, R. G., et al. 2009, *ApJ*, 695, 101
- Damjanov, I., Abraham, R. G., Glazebrook, K., et al. 2011, *ApJL*, 739, L44

- Damjanov, I., Geller, M. J., Zahid, H. J., & Hwang, H. S. 2015, *ApJ*, 806, 158
- Damjanov, I., Zahid, H. J., Geller, M. J., & Hwang, H. S. 2015, *ApJ*, 815, 104
- Davis, T. A., & McDermid, R. M. 2017, *MNRAS*, 464, 453
- de Boer, T. J. L., Belokurov, V., & Koposov, S. 2015, *MNRAS*, 451, 3489
- de Naray R. K., Spekkens K., 2011, *ApJ*, 741, L29
- Deason A. J., Auger M. W., Belokurov V., Evans N. W., 2013, *ApJ*, 773, 7
- Delaye, L., Huertas-Company, M., Mei, S., et al. 2014, *MNRAS*, 441, 203
- De Lucia, G., Springel, V., White, S. D. M., Croton, D., & Kauffmann, G. 2006, *MNRAS*, 366, 499
- Di Criscienzo, M., Caputo, F., Marconi, M., & Musella, I. 2006, *MNRAS*, 365, 1357
- Diaferio, A. 1999, *MNRAS*, 309, 610
- Di Cintio, A., Brook, C. B., Dutton, A. A., et al. 2014, *MNRAS*, 441, 2986
- Djorgovski S., Davis M., 1987, *ApJ*, 313, 59
- Dressler A., Lynden-Bell D., Burstein D., Davies R. L., Faber S. M., Terlevich R., Wegner G., 1987, *ApJ*, 313, 42
- Duffy, A. R., Schaye, J., Kay, S. T., et al. 2010, *MNRAS*, 405, 2161
- Durrell, P. R., Côté, P., Peng, E. W., et al. 2014, *ApJ*, 794, 103
- Dutton, A. A., Brewer, B. J., Marshall, P. J., et al. 2011, *MNRAS*, 417, 1621
- Ebeling, H., Voges, W., Bohringer, H., et al. 1996, *MNRAS*, 281, 799
- Ebeling H., Barrett E., Donovan D., Ma C.-J., Edge A. C., van Speybroeck L., 2007, *ApJ*, 661
- Einstein, A. 1936, *Science*, 84, 506
- Ellis R. S., Smail I., Dressler A., Couch W. J., Oemler Jr. A., Butcher H., Sharples R. M., 1997, *ApJ*, 483, 582
- Elmegreen, D. M., Chromey, F. R., Knowles, B. D., & Wittenmyer, R. A. 1998, *AJ*, 115, 1433

- El-Zant, A., Shlosman, I., & Hoffman, Y. 2001, *ApJ*, 560, 636
- El-Zant, A. A., Hoffman, Y., Primack, J., Combes, F., & Shlosman, I. 2004, *ApJL*, 607, L75
- Emsellem E., Cappellari M., Peletier R. F., McDermid R. M., Bacon R., Bureau M., Copin Y., Davies R. L., Krajnovic D., Kuntschner H., Miller B. W., de Zeeuw P. T., 2004, *MNRAS*, 352, 721
- Emsellem, E., Krajnović, D., & Sarzi, M. 2014, *MNRAS*, 445, L79
- Faber, S. M., & Jackson, R. E. 1976, *ApJ*, 204, 668
- Faber, S., Phillips, A., Kilbrick, R., et al. 2003, *SPIE*, 4841, 1657
- Fabricant, D., Fata, R., Roll, J., et al. 2005, *PASP*, 117, 1411
- Fan, L., Lapi, A., De Zotti, G., & Danese, L. 2008, *ApJL*, 689, L101
- Fan, L., Lapi, A., Bressan, A., et al. 2010, *ApJ*, 718, 1460
- Ferrarese, L., Mould, J. R., Kennicutt, R. C., Jr., et al. 2000, *ApJ*, 529, 745
- Ferrarese, L., & Merritt, D. 2000, *ApJL*, 539, L9
- Ferrarese L., et al. 2012, *ApJS*, 200, 4
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *Publ. Astron. Soc. Pacific*, 125, 306
- Forte, J. C., Martinez, R. E., & Muzzio, J. C. 1982, *AJ*, 87, 1465
- Franx, M., Illingworth, G., & Heckman, T. 1989, *AJ*, 98, 538
- Frei, Z., & Gunn, J. E. 1994, *AJ*, 108, 1476
- Gavazzi, R., Treu, T., Koopmans, L. V. E., et al. 2008, *ApJ*, 677, 1046-1059
- Gavazzi, R., Treu, T., Marshall, P. J., Brault, F., & Ruff, A. 2012, *ApJ*, 761, 170
- Gebhardt K., Kissler-Patig M., 1999, *AJ*, 118, 1526
- Geisler D., Lee M. G., Kim E., 1996, *AJ*, 111, 1529
- Gebhardt K., Adams J., Richstone D., Lauer T. R., Faber S. M., Gültekin K., Murphy J., Tremaine S., 2011, *ApJ*, 729, 119

- Gebhardt K., Thomas J., 2009, *ApJ*, 700, 1690
- Gnedin, O. Y., Kravtsov, A. V., Klypin, A. A., & Nagai, D. 2004, *ApJ*, 616, 16
- Governato F., Zolotov A., Pontzen A., Christensen C., Oh S. H., Brooks A. M., Quinn T., Shen S., Wadsley J., 2012, *MNRAS*, 422, 1231
- Graves, G. J., & Faber, S. M. 2010, *ApJ*, 717, 803
- Grillo, C. 2012, *ApJL*, 747, L15
- Hanes D., 1977, *MNRAS*, 180, 309
- Hanuschik, R. W. 2003, *A&A*, 407, 1157
- Harris W. E., 2009, *ApJ*, 703, 939
- Harris, W. E. 1991, *ARA&A*, 29, 543
- Harris W. E., Morningstar W., Gnedin O. Y., O'Halloran H., Blakeslee J. P., Whitmore B. C., Cote P., Geisler D., Peng E. W., Bailin J., Rothberg B., Cockcroft R., DeGraaff R. B., 2014, *ApJ*, 797, 128
- Harvey, D., Courbin, F., Kneib, J. P., & McCarthy, I. G. 2017, [arXiv:1703.07365](https://arxiv.org/abs/1703.07365)
- Hausman, M. A., & Ostriker, J. P. 1978, *ApJ*, 224, 320
- Hilz, M., Naab, T., & Ostriker, J. P. 2013, *MNRAS*, 429, 2924
- Hogg, D. W., Baldry, I. K., Blanton, M. R., & Eisenstein, D. J. 2002, [arXiv:astro-ph/0210394](https://arxiv.org/abs/astro-ph/0210394)
- Hogg D. W., Bovy J., Lang D., 2010, *ArXiv e-prints*
- Holden B. P., van der Wel A., Kelson D. D., Franx M., Illingworth G. D., 2010, *ApJ*, 724, 714
- Hopkins, P. F., Bundy, K., Murray, N., et al. 2009, *MNRAS*, 398, 898
- Hopkins, P. F., Bundy, K., Hernquist, L., Wuyts, S., & Cox, T. J. 2010, *MNRAS*, 401, 1099
- Host, O., & Hansen, S. H. 2011, *ApJ*, 736, 52
- Houghton R. C. W., Davies R. L., Dalla Bontà E., Masters R., 2012, *MNRAS*, 423, 256
- Hsu, L.-Y., Stockton, A., & Shih, H.-Y. 2014, *ApJ*, 796, 92

- Huertas-Company, M., Shankar, F., Mei, S., Bernardi, M., Aguerri, J. A. L., Meert, A., Vikram, V., 2013, *ApJ*, 779, 29
- Hubble, E. P. 1926, *ApJ*, 64,
- Hubble, E. P. 1927, *The Observatory*, 50, 276
- Hunter D. A., Brinks E., Elmegreen B., Rupen M., Simpson C., Walter F., Westpfahl D., Young L., 2007, *Am. Astron. Soc.*, 39
- Hyde, J. B., & Bernardi, M. 2009, *MNRAS*, 396, 1171
- Ishibashi, W., Fabian, A. C., & Canning, R. E. A. 2013, *MNRAS*, 431, 2350
- Jørgensen I., Franx M., Kjaergaard P., 1995, *MNRAS*, 273, 1097
- Jorgensen, I., Franx, M., & Kjaergaard, P. 1996, *MNRAS*, 280, 167
- Jørgensen I., Franx M., Hjorth J., van Dokkum P. G., 1999a, *MNRAS*, 308, 833
- Jørgensen I., Chiboucas K., Flint K., Bergmann M., Barr J., Davies R., 2006, *ApJ*, 639
- Jørgensen, I., & Chiboucas, K. 2013, *AJ*, 145, 77
- Jørgensen, I., Chiboucas, K., Toft, S., et al. 2014, *AJ*, 148, 117
- Kelly B. C., 2007, *ApJ*, 665, 1489
- Kim S., Rey S.-C., Jerjen H., Lisker T., Sung E.-C., Lee Y., Chung J., Pak M., Yi W., Lee W., 2014, *ApJSuppl. Ser.*, 215, 22
- Klypin A., Holtzman J., 1997, eprint arXiv:astro-ph/9712217
- Ko, Y., Hwang, H. S., Lee, M. G., et al. 2017, *ApJ*, 835, 212
- Kobayashi, C. 2004, *PASA*, 21, 183
- Kodama T., Arimoto N., 1997, *Å*, 320, 41
- Koopmans, L. V. E., & Treu, T. 2003, *ApJ*, 583, 606
- Koopmans, L. V. E. 2005, *MNRAS*, 363, 1136
- Koopmans, L. V. E., Treu, T., Bolton, A. S., Burles, S., & Moustakas, L. A. 2006, *ApJ*, 649, 599

- Koopmans, L. V. E., Bolton, A., Treu, T., et al. 2009, *ApJL*, 703, L51
- Kormendy, J. 1977, *ApJ*, 218, 333
- Kormendy J., Fisher D. B., Cornell M. E., Bender R., 2009, *ApJ*, 182, 216
- Kroupa, P. 2001, *MNRAS*, 322, 231
- Kuntschner, H., Emsellem, E., Bacon, R., et al. 2010, *MNRAS*, 408, 97
- La Barbera, F., Busarello, G., Merluzzi, P., et al. 2008, *ApJ*, 689, 913-918
- Lagattuta, D. J., Auger, M. W., & Fassnacht, C. D. 2010, *ApJL*, 716, L185
- Lani, C., Almaini, O., Hartley, W. G., et al. 2013, *MNRAS*, 435, 207
- Laporte C. F. P., White S. D. M., Naab T., Ruszkowski M., Springel V., 2012, *MNRAS*, 424, 747
- Laporte C. F. P., & White S. D. M. 2015, *MNRAS*, 451, 1177
- Larson, R. B. 1990, *PASP*, 102, 709
- Macciò, A. V., Dutton, A. A., & van den Bosch, F. C. 2008, *MNRAS*, 391, 1940
- Mamon G. A., Łokas E. L., 2005, *MNRAS*, 363, 705
- Marshall, P. J., Treu, T., Melbourne, J., et al. 2007, *ApJ*, 671, 1196
- Martín-Navarro, I., Barbera, F. L., Vazdekis, A., Falcón-Barroso, J., & Ferreras, I. 2015, *MNRAS*, 447, 1033
- Martín-Navarro, I., Vazdekis, A., La Barbera, F., et al. 2015, *ApJL*, 806, L31
- Martizzi D., Teyssier R., Moore B., Wentz T., 2012, *MNRAS*, 422, 3081
- Martizzi, D., Teyssier, R., & Moore, B. 2013, *MNRAS*, 432, 1947
- Mashchenko S., Couchman H. M. P., Wadsley J., 2006, *Nature*, 442, 539
- Massey, R., Williams, L., Smit, R., et al. 2015, *MNRAS*, 449, 3393
- McConnell, N. J., & Ma, C.-P. 2013, *ApJ*, 764, 184
- McKee, C. F., & Ostriker, E. C. 2007, *ARA&A*, 45, 565

- McLaughlin D. E., Harris W. E., Hanes D. A., 1994, *ApJ*, 422, 486
- Mead, J. M. G., King, L. J., Sijacki, D., et al. 2010, *MNRAS*, 406, 434
- Mei, S., Blakeslee, J. P., Côté, P., et al. 2007, *ApJ*, 655, 144
- Mei, S., Holden, B. P., Blakeslee, J. P., et al. 2009, *ApJ*, 690, 42
- Merritt, D. 1985, *AJ*, 90, 1027
- Montes M., Trujillo I., Prieto M. A., Acosta-Pulido J. A., 2014, *MNRAS*, 439, 990
- Morishita, T., Abramson, L. E., Treu, T., et al. 2016, arXiv:1607.00384
- Moster, B. P., Somerville, R. S., Maulbetsch, C., et al. 2010, *ApJ*, 710, 903
- Muñoz, J. A., Kochanek, C. S., & Keeton, C. R. 2001, *ApJ*, 558, 657
- Murphy J. D., Gebhardt K., Adams J. J., 2011, p. 29
- Naab, T., Johansson, P. H., & Ostriker, J. P. 2009, *ApJL*, 699, L178
- Napolitano N. R., Romanowsky A. J., & Tortora C. 2010, *MNRAS*, 405, 2351
- Napolitano N. R., Pota V., Romanowsky A. J., Forbes D. A., Brodie J. P., Foster C., 2014, *MNRAS*, 439, 659
- Navarro J. F., Eke V. R., & Frenk C. S. 1996, *MNRAS*, 283, L72
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
- Navarro, J. F., Ludlow, A., Springel, V., et al. 2010, *MNRAS*, 402, 21
- Newman, A. B., Treu, T., Ellis, R. S., et al. 2009, *ApJ*, 706, 1078
- Newman A. B., Treu T., Ellis R. S., Sand D. J., 2011, *ApJ*, 728, L39
- Newman, A. B., Ellis, R. S., Bundy, K., & Treu, T. 2012, *ApJ*, 746, 162
- Newman, A. B., Treu, T., Ellis, R. S., & Sand, D. J. 2013, *ApJ*, 765, 25
- Newman, A. B., Ellis, R. S., Andreon, S., et al. 2014, *ApJ*, 788, 51
- Newman, A. B., Ellis, R. S., & Treu, T. 2015, *ApJ*, 814, 26
- Newton, E. R., Marshall, P. J., Treu, T., et al. 2011, *ApJ*, 734, 104

- Nipoti, C., Treu, T., Ciotti, L., & Stiavelli, M. 2004, MNRAS, 355, 1119
- Nipoti, C., Treu, T., Leauthaud, A., et al. 2012, MNRAS, 422, 1714
- Oke, J., Cohen, J., Carr, M., et al. 1995, PASP, 107, 375
- Oldham L. J., & Auger M. W. 2016, MNRAS, 455, 820
- Osipkov, L. P. 1979, Pisma v Astronomicheskii Zhurnal, 5, 77
- Ostrov P., Geisler D., Forte J. C., 1993, AJ, 105, 1762
- O’Sullivan, E., Forbes, D. A., & Ponman, T. J. 2001, MNRAS, 328, 461
- Peletier, R. F., Davies, R. L., Illingworth, G. D., Davis, L. E., & Cawson, M. 1990, AJ, 100, 1091
- Peng E. W., Jordán A., Blakeslee J. P., Mieske S., Côté P., Ferrarese L., Harris W. E., Madrid J. P., Meurer G. R., 2009, ApJ, 703, 42
- Peralta de Arriba, L., Balcells, M., Falcón-Barroso, J., & Trujillo, I. 2014, MNRAS, 440, 1634
- Poggianti, B. M., Calvi, R., Bindoni, D., et al. 2013, ApJ, 762, 77
- Pontzen A., & Governato F. 2012, MNRAS, 421, 3464
- Posti, L., Nipoti, C., Stiavelli, M., & Ciotti, L. 2014, MNRAS, 440, 610
- Postman, M., Coe, D., Benítez, N., et al. 2012, ApJS, 199, 25
- Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, MNRAS, 423, 3018
- Read, J. I., & Gilmore, G. 2005, MNRAS, 356, 107
- Saglia R. P., Sanchez-Blazquez P., Bender R., Simard L., Desai V., Aragon-Salamanca A., Milvang-Jensen B., Halliday C. e. a., 2010, A& A, 524
- Salpeter, E. E. 1955, ApJ, 121, 161
- Sand, D. J., Treu, T., & Ellis, R. S. 2002, ApJL, 574, L129
- Sand D. J., Treu T., Smith G. P., Ellis R. S., 2004, ApJ, 604, 88

- Sand, D. J., Treu, T., Ellis, R. S., Smith, G. P., & Kneib, J.-P. 2008, *ApJ*, 674, 711
- Sandage A., 1972, *ApJ*, 176, 21
- Saracco, P., Casati, A., Gargiulo, A., et al. 2014, *A&A*, 567, A94
- Sarzi, M., Spiniello, C., et al., submitted
- Schaller, M., Frenk, C. S., Bower, R. G., et al. 2015, *MNRAS*, 452, 343
- Schlafly, E. F., & Finkbeiner, D. P. 2011, *ApJ*, 737, 103
- Schneider, P., Ehlers, J., & Falco, E. E. 1992, *Gravitational Lenses*, XIV, 560 pp. 112 figs.. Springer-Verlag Berlin Heidelberg New York. Also *Astronomy and Astrophysics Library*, 112
- Schuberth Y., Richtler T., Hilker M., Dirsch B., Bassino L. P., Romanowsky A. J., Infante L., 2010, *A&AS*, 513, A52
- Shankar, F., Marulli, F., Bernardi, M., et al. 2013, *MNRAS*, 428, 109
- Sharma S., Bland-Hawthorn J., Johnston K. V., Binney J., 2011, *ApJ*, 730, 3
- Sheinis, A. I., Bolte, M., Epps, H. W., et al. 2002, *PASP*, 114, 851
- Shen, S., Mo, H. J., White, S. D. M., et al. 2003, *MNRAS*, 343, 978
- Simard, L., Mendel, J. T., Patton, D. R., Ellison, S. L., & McConnachie, A. W. 2011, *ApJS*, 196, 11
- Smith, R. J., Lucey, J. R., & Conroy, C. 2015, *MNRAS*, 449, 3441
- Sonnenfeld, A., Treu, T., Gavazzi, R., et al. 2012, *ApJ*, 752, 163
- Sonnenfeld, A., Treu, T., Gavazzi, R., et al. 2013, *ApJ*, 777, 98
- Sonnenfeld, A., Nipoti, C., & Treu, T. 2014, *ApJ*, 786, 89
- Sonnenfeld, A., Treu, T., Marshall, P. J., et al. 2015, *ApJ*, 800, 94
- Stanford S. A., Eisenhardt P. R., Dickinson M., 1995, *ApJ*, 450, 512
- Stanford S. A., Eisenhardt P. R., Dickinson M., 1998, *ApJ*, 492, 461
- Stockton, A., Shih, H.-Y., & Larson, K. 2010, *ApJL*, 709, L58

- Stockton, A., Shih, H.-Y., Larson, K., & Mann, A. W. 2014, *ApJ*, 780, 134
- Stott, J. P., Collins, C. A., Burke, C., Hamilton-Morris, V., & Smith, G. P. 2011, *MNRAS*, 414, 445
- Strader J., Brodie J., Spitler L., Beasley M., 2006, *AJ*, 132, 2333
- Strader J. et al., 2011, *ApJS*, 197, 33
- Strauss, M. A., Weinberg, D. H., Lupton, R. H., et al. 2002, *AJ*, 124, 1810
- Stringer, M., Trujillo, I., Dalla Vecchia, C., & Martinez-Valpuesta, I. 2015, *MNRAS*, 449, 2396
- Struck, C., & Smith, B. J. 2012, *MNRAS*, 422, 2444
- Tamura, N., & Ohta, K. 2003, *AJ*, 126, 596
- Tamura N., Sharples R. M., Arimoto N., Onodera M., Ohta K., Yamada Y., 2006a, *MNRAS*, 373, 588
- Tamura N., Sharples R. M., Arimoto N., Onodera M., Ohta K., Yamada Y., 2006b, *MNRAS*, 373, 601
- Tanaka, M., Wong, K. C., More, A., et al. 2016, *ApJL*, 826, L19
- Taylor, E. N., Franx, M., Glazebrook, K., et al. 2010, *ApJ*, 720, 723
- Terlevich A. I., Caldwell N., Bower R. G., 2001, *MNRAS*, 326, 1547
- Tody, D. 1993, *Astronomical Data Analysis Software and Systems II*, 52, 173
- Tortora, C., Antonuccio-Delogu, V., Kaviraj, S., et al. 2009, *MNRAS*, 396, 61
- Tortora, C., Napolitano, N. R., Cardone, V. F., et al. 2010, *MNRAS*, 407, 144
- Tortora, C., Pipino, A., D’Ercole, A., Napolitano, N. R., & Matteucci, F. 2013, *MNRAS*, 435, 786
- Tortora, C., La Barbera, F., Napolitano, N. R., et al. 2016, *MNRAS*, 457, 2845
- Treu, T., & Koopmans, L. V. E. 2002, *ApJ*, 575, 87
- Treu, T., & Koopmans, L. V. E. 2004, *ApJ*, 611, 739

- Treu, T., Ellis, R. S., Liao, T. X., et al. 2005, *ApJ*, 633, 174
- Treu, T., Auger, M. W., Koopmans, L. V. E., et al. 2010, *ApJ*, 709, 1195
- Treu, T. 2010, *ARA&A*, 48, 87
- Trujillo, I., Burkert, A., & Bell, E. F. 2004, *ApJL*, 600, L39
- Trujillo, I., Feulner, G., Goranova, Y., et al. 2006, *MNRAS*, 373, L36
- Trujillo I., Conselice C. J., Bundy K., Cooper M. C., Eisenhardt P., Ellis R. S., 2007, *MNRAS*, 382, 109
- Valdes F., Gupta R., Rose J. A., Singh H. P., Bell D. J., 2004, *ApJS*, 152, 251
- Valentinuzzi, T., Fritz, J., Poggianti, B. M., et al. 2010, *ApJ*, 712, 226
- Valentinuzzi, T., Poggianti, B. M., Saglia, R. P., et al. 2010, *ApJL*, 721, L19
- van der Marel R. P., 1994, *MNRAS*, 270
- van der Wel A., Holden B. P., Zirm A. W., Franx M., Rettura A., Illingworth G. D., Ford H. C., 2008, *ApJ*, 688, 48
- van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, *ApJ*, 788, 28
- van Dokkum P. G., Franx M., 1996, *MNRAS*, 281, 985
- van Dokkum, P. G., & Franx, M. 2001, *ApJ*, 553, 90
- van Dokkum P. G., Ellis R. S., 2003, *ApJ*, 592
- van Dokkum, P. G., Franx, M., Kriek, M., et al. 2008, *ApJL*, 677, L5
- van Dokkum, P. G., & Conroy, C. 2010, *Nature*, 468, 940
- van Dokkum, P. G., Whitaker, K. E., Brammer, G., et al. 2010, *ApJ*, 709, 1018
- van Dokkum, P. G., Nelson, E. J., Franx, M., et al. 2015, *ApJ*, 813, 23
- van Dokkum, P., Conroy, C., Villaume, A., Brodie, J., & Romanowsky, A. 2016, [arXiv:1611.09859](https://arxiv.org/abs/1611.09859)
- Vaughan, S. P., Houghton, R. C. W., Davies, R. L., & Zieleniewski, S. 2016, [arXiv:1612.00364](https://arxiv.org/abs/1612.00364)
- Vegetti, S., & Koopmans, L. V. E. 2009, *MNRAS*, 392, 945

- Velliscig M., van Daalen M. P., Schaye J., McCarthy I. G., Cacciato M., Le Brun A. M. C., Vecchia C. D., 2014, MNRAS, 442, 2641
- Velliscig, M., Cacciato, M., Schaye, J., et al. 2015, MNRAS, 453, 721
- Visvanathan N., Sandage A., 1977, ApJ, 216, 214
- Walker M. G., Peñarrubia J., 2011, ApJ, 742, 20
- Warren, S. J., & Dye, S. 2003, ApJ, 590, 673
- Watkins L. L., Evans N. W., An J. H., 2010, MNRAS, 406, 264
- White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341
- Wolf, J., Martinez, G. D., Bullock, J. S., et al. 2010, MNRAS, 406, 1220
- Wojtak, R., Łokas, E. L., Gottlöber, S., & Mamon, G. A. 2005, MNRAS, 361, L1
- Wu, X., Gerhard, O., Naab, T., et al. 2014, MNRAS, 438, 2701
- Wuyts, S., Cox, T. J., Hayward, C. C., et al. 2010, ApJ, 722, 1666
- Zemp, M., Gnedin, O. Y., Gnedin, N. Y., & Kravtsov, A. V. 2012, ApJ, 748, 54
- Zepf S., Ashman K., 1993, MNRAS, 264
- Zhu, L., et al. 2014, ApJ, 792, 59
- Zieleniewski, S., Houghton, R. C. W., Thatte, N., Davies, R. L., & Vaughan, S. P. 2017, MNRAS, 465, 192